Docket No.: 28341/6223.NDV1

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
APPLICATION FOR UNITED STATES LETTERS PATENT

Title:

G PROTEIN-COUPLED RECEPTOR-LIKE RECEPTORS AND MODULATORS THEREOF

David E. Lowery	1207 Woodland Drive Portage, Michigan 49024 Citizenship: U.S.
Timothy G. Geary	5176 Burning Tree Kalamazoo, Michigan 49009 Citizenship: U.S.
Teresa M. Kubiak	5844 East B. Avenue Richland, Michigan 49083 Citizenship: Poland
Martha J. Larsen	56 Naples Court Kalamazoo, Michigan 49009 Citizenship: U.S.

G Protein-Coupled Receptor-Like Receptors and Modulators Thereof

FIELD OF THE INVENTION

The present invention relates generally to G-protein, seven-transmembrane-receptor-like polynucleotides and polypeptides, as well as modulators of the interaction between the polypeptides and their ligand(s) that are useful for inhibiting the neurological activity of organisms, *e.g.*, invertebrates.

BACKGROUND

It has been estimated that as many as one in three humans is infected by one or more species of parasitic helminth (flatworms and roundworms). Helminths also represent a chronic and continuing threat to the health of livestock and companion animals; their abundance and resistance to anthelmintic drugs effectively prevents animal agriculture in certain environments. Despite the prevalence of helminth-caused disease, little is known of helminth physiology. This ignorance, in turn, hampers efforts to identify agents capable of controlling these pathogens.

Flatworms (platyhelminths) are evolutionarily quite distinct from the roundworms (nematodes) and differ markedly in neuromuscular anatomy and physiology. Only a subset of one class of drugs (some of the benzimidazoles) shows activity against both flatworms and roundworms. The diversity among the animal species that threaten the health of man, livestock, crops, and sensitive environmental niches, presents a challenge to efforts to mount broad-based attacks on the pest organisms.

G protein-coupled receptors (*i.e.*, GPCRs) form a vast superfamily of cell surface receptors which are present in virtually all animal cells and are characterized by an amino-terminal extracellular domain, a carboxy-terminal intracellular domain, and a serpentine structure that passes through the cell membrane seven times. Hence, such receptors are sometimes also referred to as seven transmembrane (7TM) receptors. These seven transmembrane domains define three extracellular loops and three intracellular loops, in addition to the amino- and carboxy-terminal domains. The extracellular portions of the receptor have a role in recognizing and binding one or more extracellular binding partners (*e.g.*, ligands), whereas the intracellular portions have a role in recognizing and communicating with downstream effector molecules.

The G protein-coupled receptors bind a variety of ligands including calcium ions, hormones, chemokines, neuropeptides, neurotransmitters, nucleotides, lipids, odorants, and even photons. Not surprisingly, the GPCRs are important in the normal (and sometimes the aberrant) function of many cell types. [See generally Strosberg, Eur. J. Biochem., 196: 1-10 (1991) and Bohm *et al.*, Biochem J., 322: 1-18 (1997).] When a specific ligand binds to its corresponding receptor, the ligand typically stimulates the receptor to activate a specific heterotrimeric guanine nucleotide-binding regulatory protein (G protein) that is coupled to the intracellular portion or region of the receptor. The G protein, in turn, transmits a signal to an effector molecule within the cell by either stimulating or inhibiting the activity of that effector molecule. These effector molecules include adenylate cyclase, phospholipases and ion channels. Adenylate cyclase and phospholipases are enzymes that are involved in the production of the second messenger molecules cAMP, inositol triphosphate and diacyglycerol. It is through this sequence of events that an extracellular ligand stimulus exerts intracellular changes through a G protein-coupled receptor. Each such receptor has its own characteristic primary structure, expression pattern, ligand binding profile, and intracellular effector system.

Because of the vital role of G protein-coupled receptors in the communication between cells and their environment, such receptors are attractive targets for therapeutic intervention, and drugs that activate or antagonize the activation of such receptors are known. For receptors having a known ligand, the identification of agonists or antagonists may be sought specifically for mimicking, enhancing or inhibiting the action of the ligand. Thus, GPCRs show promise as potential targets of methods for treating infestations and/or infections caused by a variety of invertebrate pests, including both ecto- and endo-parasites. However, such methods must be able to discriminate the GPCRs of the invertebrate pest organisms from the GPCRs found in those species of plants and vertebrate animals upon whom the pests prey.

A large family of peptides (typically 4-15 amino acids in length) that is largely, if not exclusively, found in invertebrate animals such as helminths is a class of neuropeptides known as FMRFamide related peptides (*i.e.*, FaRPs). The prototypical FMRFamide peptides are so named because of the "FMRF" amino acid sequence, including

the consensus "RF" sequence, at their C-termini. As neuropeptides, these molecules are involved in vital biological processes requiring controlled neuromuscular activity. Although some neurotransmitters and neuromodulators (including neuropeptides) have been shown to function as ligands for receptors, to date there has been no identification of a FaRP neuropeptide as a ligand of a GPCR.

Because of the toxic potential of broad-spectrum chemical parasiticides, there exists a need in the art for targeted biologicals capable of selectively interfering with the life cycle of harmful invertebrates such as helminths and insects without harming host plant and animal species, as well as the environment.

SUMMARY OF THE INVENTION

The present invention generally relates to materials and methods for the targeted interference with vital biological processes of pest invertebrates. By providing materials and methods for modulating the activity of invertebrate GPCRs involved in neuromuscular activity, the invention provides a biological approach to invertebrate pest control that can minimize the deleterious consequences to non-pest species of animals, including man, as well as plants and the environment in general.

One aspect of the invention is a screening method for identifying candidate anti-invertebrate modulators that affect one or more activities of an invertebrate GPCR-like receptor involved in neuromuscular functioning including, *e.g.*, binding of a GPCR-like receptor to a ligand, typically a peptide ligand, and signal transduction. The method comprises the steps of: (a) contacting a test compound with a composition, wherein the composition contains a GPCR-like receptor encoded by a polynucleotide having a sequence selected from the group consisting of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176 and 178, or a polynucleotide hybridizing to the GPCR-like receptor under stringent conditions of hybridizing at 42°C in a solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS; and (b) measuring the activity of the GPCR-like receptor in the presence and absence of the test compound. In preferred embodiments, the screening method includes a receptor, or fragment, variant or derivative thereof, having a sequence set forth in SEQ ID NOS: 7, 21, 25, 35, 43, 45, 105, 107, 109, 111, 113, 115, 117, 177 and 179.

As one of ordinary skill in the art would recognize, the above-described method can be practiced with a variety of GPCR-like receptors. For example, the GPCR-like receptor used in the screening method may be encoded by a polynucleotide having a sequence set forth in any one of SEQ ID NOS:43, 21, 45, 35, 7, 106, and 104. As noted above, such GPCR-like receptors may be used in screening assays designed to measure a GPCR-like receptor activity, including binding activity. Expressly contemplated are embodiments of the screening method comprising a GPCR-like receptor encoded by a polynucleotide comprising a sequence set forth in SEQ ID NO:43 and a peptide comprising

a sequence selected from the group consisting of SEQ ID NOS:85, 86, 88, 89, and 118, wherein the peptide binds to the GPCR-like receptor.

encoded by a polynucleotide having the sequence set forth in SEQ ID NO:43 is provided. In an alternative embodiment, the GPCR-like receptor comprises a sequence set forth in SEQ ID NO:21 and the peptide comprises a sequence selected from the group consisting of SEQ ID NOS:78, 79, 80, 84, 87, 92, 98, 100, 120, 171, 143, 122, 123, 97, 85, 83, 101, 102, 93, 88, 91, 94, 93, 90, 152, 153, 154, 155, 156, 157, 80, 158, 119, 159, 160, 161, 162, 163 and 164. Another embodiment involves a GPCR-like receptor comprising a sequence set forth in SEQ ID NO:45 and a peptide comprising a sequence selected from the group consisting of SEQ ID NOS:86, 118, 125, 88, 126, 127, 128, 129, 102, 131, 100, 133, 92, 135, 136, 137, 87, 139, 91, 141 and 83. In yet another embodiment, the GPCR-like receptor comprises a sequence set forth in SEQ ID NO:35 and a peptide comprising a sequence selected from the group consisting of SEQ ID NOS:99, 97, 96, 77, 82, 81, 87, 100, 92, 80, 98, 120, 121, 79 and 84. In still another embodiment, the GPCR-like receptor comprises a sequence set forth in SEQ ID NO:7 and a peptide comprising a sequence selected from the group consisting of SEQ ID NOS:94, 103, 95, 101, 85, 79, 84, 87, 86, 80, 92, 100, and 180. Yet another embodiment involves a GPCR-like receptor comprising a sequence selected from the group consisting of SEQ ID NO:106 and a peptide comprising a sequence selected from the group consisting of SEQ ID NOS:80, 92, 98, 100, 120, 121, 79, 84, 136, 87 and 86. Still another one of the many embodiments of this aspect of the invention involves a GPCR-like receptor comprising a sequence set forth in SEQ ID NO:104 and a peptide comprising a sequence selected from the group consisting of SEQ ID NOS:80, 92, 98, 100, 120, 121, 79, 84, 136, 87, 86, 150, 151, 133, 165, 91, 166, 131 and 167.

Another aspect of the invention is a method of identifying an anti-invertebrate modulator of an activity of an invertebrate GPCR-like receptor comprising the following steps: (a) contacting a test compound and a composition, wherein the composition contains a GPCR-like receptor selected from the group consisting of polypeptides encoded by a polynucleotide having a sequence selected from the group consisting of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176 and 178, or a polynucleotide hybridizing to the GPCR-like receptor

under stringent conditions of hybridizing at 42°C in a solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% Dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS, optionally in the presence of a peptide or other ligand of the receptor; and (b) measuring the activity of the GPCR-like receptor in the presence and absence of the test compound. Modulators are identified as test compounds that alter (*i.e.*, increase or decrease) a GPCR-like receptor function, such as a binding property of a receptor or an activity such as G protein-mediated signal transduction or membrane localization. The composition may contain an isolated GPCR-like receptor; alternatively, the composition may contain a GPCR-like receptor in association with, *e.g.*, an intact cell or cell portion, such as a membrane. Presently preferred embodiments of the method use a GPCR-like receptor having an amino acid sequence set forth in SEQ ID NOS: 8, 22, 26, 36, 44, 105, 107, 109, 111, 113, 115, 117, 177 or 179. Preferred peptides are neuropeptides derived from invertebrates and include the FaRP family of neuropeptides. Particularly preferred are invertebrate neuropeptides having an amino acid sequence selected from the group consisting of SEQ ID NOS: 77-103 and 118-151. The methods of the invention embrace neuropeptides that are attached to a label, such as a radiolabel (*e.g.*, ^{125}I, ^{35}S, ^{32}P, ^{33}P, ^{3}H), a fluorescence label, a chemiluminescence label, an enzymic label and an immunogenic label.

In various embodiments of the method, the assay may take the form of an ion flux assay, a yeast growth assay, a non-hydrolyzable GTP assay such as a [^{35}S]-GTPγS assay, a cAMP assay, an inositol triphosphate assay, a diacylglycerol assay, an Aequorin assay, a Luciferase assay, a FLIPR assay for intracellular Ca^{2+} concentration, a mitogenesis assay, a MAP Kinase activity assay, an arachidonic acid release assay (*e.g.*, using [^{3}H]-arachidonic acid), and an assay for extracellular acidification rates, as well as other binding or function-based assays of GPCR activity that are generally known in the art. In several of these embodiments, the invention comprehends the inclusion of any of the G proteins known in the art, such as $G_{\alpha 16}$, $G_{\alpha 15}$, or chimeric G_{qi5}, G_{qs5}, G_{qo5}, or G_{qz5}.

In another aspect of the invention, a method of identifying a candidate anti-invertebrate modulator is provided. The method comprises the steps of: (a) contacting a test compound and a composition, wherein the composition contains a GPCR-like receptor

encoded by a polynucleotide selected from the group consisting of receptor polynucleotides having a sequence set forth in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176 and 178, and polynucleotides hybridizing to the receptor polynucleotides under stringent conditions of hybridizing at 42°C in a solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS ; and (b) identifying a test compound that binds to or interacts with the composition as a candidate anti-invertebrate modulator. In a preferred embodiment, the polynucleotide encoding the GPCR-like receptor comprises a sequence selected from the group consisting of SEQ ID NOS: 7, 21, 25, 35, 43, 45, 105, 107, 109, 111, 113, 115, 117, 177 and 179.

Another aspect of the invention is a method of identifying an anti-invertebrate agent comprising the following steps: (a) identifying a modulator using the method described above; (b) contacting the modulator and an invertebrate tissue; and (c) measuring the response of the invertebrate tissue, *e.g.* neural signaling or neuromuscular activity, thereby identifying the modulator as an anti-invertebrate agent. Although any invertebrate tissue may be used in the method, presently preferred tissue sources are helminths and insects. An exemplary type of suitable invertebrate tissue is neuromuscular tissue, e.g. in isolated form or remaining in association with part, or all, of an invertebrate organism. The anti-invertebrate agents, and compositions comprising one of those agents, identified by this method constitute yet another aspect of the invention.

In a related aspect, the invention provides a method for treating an invertebrate comprising the step of contacting the invertebrate with a biologically effective amount of a modulator identified by the methods described herein. A biologically effective amount of a modulator is an amount that is sufficient to induce a desired response in the treated invertebrate. Thus, a biologically effective amount of a modulator may be the amount that interferes with physiological activity of the treated invertebrate in a non-lethal manner (*i.e.*, a biostatic effect) or in a lethal manner (*i.e.*, a biocidal effect). A preferred modulator for use in the treatment methods is an inhibitor of GPCR-like receptor activity. The invention is not limited to particular means for delivering the modulator to an

invertebrate, nor is the invention limited as to the compositions comprising the modulator which may be delivered.

Another aspect of the invention is drawn to methods of producing an invertebrate GPCR-like receptor comprising the following steps:(a) incubating a source cell at a temperature of at least about 35°C; (b) lowering the temperature to at most about 26°C; and (c) detecting the GPCR-like receptor. In various embodiments, the temperatures may be varied to optimize production using no more than routine experimentation. For example, the cells may be incubated at temperatures higher than about 35°C, *e.g.*, a temperature of at least about 37°C; the temperature also may be lowered beyond 30°C, for example to at most 29°C, at most 28°C, at most 27°C, at most 26°C, at most 25°C or at most 24°C. In some embodiments, both the incubation temperature may be raised above 35°C and the temperature lowering may extend beyond 30°C. The method of producing an invertebrate GPCR-like receptor may further comprise recovering the GPCR-like receptor, which may be native or recombinant in origin. The receptor may be recovered in intact cells, cell portions (*e.g.*, membranes) obtained as a result of cell lysis, or in isolated form. Any of a wide variety of cells may be used as source cells, such as cells derived from mammals, amphibians, arthropods (*e.g.*, insects), mollusks, helminths, and others. It is anticipated that this method of producing a GPCR-like receptor is particularly suited to the recombinant production of GPCR-like receptors using non-invertebrate cells, such as mammalian cells.

The invention also comprehends compositions of matter, such as a modulator of an activity of the GPCR-like receptors identified by the methods described herein. The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into non-peptide mimetics of natural GPCR-like receptor ligands, peptide and non-peptide allosteric effectors of GPCR-like receptors, and peptides and non-peptide compounds that function as activators or inhibitors (competitive, uncompetitive and non-competitive) (*e.g.*, antibody products) of an ultimate GPCR-like receptor activity. For example, such modulators may be compounds or compositions that are agonists or antagonists of ligand binding to a GPCR-like receptor, allosteric effectors thereof, or compounds (or compositions) that affect the ultimate activity of GPCR-like receptors through direct action on the receptor or an effect introduced downstream of the receptor in,

e.g., a signal cascade. In addition, modulators according to the invention may be compounds or compositions that interfere with the expression of a GPCR-like receptor, either through inhibiting transcription of the DNA or translation of the corresponding mRNA. Expression of the GPCR-like receptor can be monitored by any methods known in the art, including Western blot analysis using polyclonal or monoclonal antibodies to the GPCR-like receptor or Northern blot analysis or quantitative polymerase chain reaction (PCR) using suitable probes or primers based on the sequence of the GPCR-like receptor gene. In particular, modulators that interfere with the expression of gene products include anti-sense polynucleotides and ribozymes that are complementary to the gene sequences. The invention further embraces modulators that affect the transcription of gene products of the invention through the formation of oligonucleotide-directed triplet helix formation. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries. Preferred peptide receptors have an amino acid sequence selected from the group consisting of SEQ ID NOS: 8, 22, 26, 36, 44, 105, 107, 109, 111, 113, 115, 117, 177 or 179.

Another aspect of the invention is drawn to an isolated GPCR-like receptor comprising a sequence selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 105, 107, 109, 111, 113, 115, 117, 177, and 179, and variants and fragments thereof. Preferred receptors have sequences set forth in SEQ ID NOS: 8, 22, 26, 36, 44, 105, 107, 109, 111, 113, 115, 117, 177 and 179. Variants and fragments of GPCR-like receptors retain at least one biological or immunological property of the cognate GPCR-like receptor, and are at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 125, at least 150, at least 175, at least 200, at least 225, at least 250, at least 275, at least 300, at least 325, at least 350, at least 375, at least 400, at least 425, or at least 450 amino acids in length. Fragments specifically include GPCR domains of the receptors. Biological activities of GPCR-like receptors according to the invention include, but are not limited to, the binding of a natural or an unnatural ligand, as

well as any one of the functional activities of GPCRs known in the art. Other non-limiting examples of GPCR activities include transmembrane signaling of various forms, which may involve G protein association and/or the exertion of an influence over G protein binding of various guanidylate nucleotides; transmembrane localization; or binding of accessory proteins or polypeptides unrelated to known G proteins.

The invention also provides an isolated polynucleotide encoding a GPCR-like receptor. Such polynucleotides may be selected from the group consisting of: (a) a polynucleotide comprising a nucleotide sequence encoding any one of the amino acid sequences set forth in SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 105, 107, 109, 111, 113, 115 117, 177, and 179 (including the nucleotide sequences set forth in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176, and 178); and (b) a polynucleotide which hybridizes under conditions of high stringency to the complement of the polynucleotide of (a). Exemplary conditions of high stringency are provided below. Such polynucleotides also include polynucleotides that exhibit at least 90%, at least 95%, at least 98%, at least 99% or at least 99.9% sequence identity to either a polynucleotide sequence disclosed in the sequence listing (*i.e.*, SEQ ID NOS: 1, 3, 5, 7, 9, 11, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176 and 178) or to a polynucleotide encoding a GPCR-like receptor comprising one of the sequences disclosed in the sequence listing (*i.e.*, SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 105, 107, 109, 111, 113, 115, 117, 177, and 179). Any one of the publicly available algorithms (*e.g.*, the BLASTl program of GCG) for comparing sequences may be used in determining the degree of sequence similarity (including appropriate penalties for gap introductions). A preferred algorithm is the BLAST algorithm implemented at the GenBank website under the auspices of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/) using default parameters. A polynucleotide of the invention may be partially or wholly chemically synthesized and embraces an anti-sense polynucleotide which specifically hybridizes to the complement of one or more of the above-described polynucleotides. In related aspects, the invention comprehends

vectors comprising these polynucleotides preferably operably linked to expression control sequences, including expression vectors, as well as non-native host cells transformed or transfected with a polynucleotide in accordance with the invention or a host cell transformed or transfected with a vector of the invention. All suitable native and non-native host cells are embraced by the invention, including mammalian cells (*e.g.*, COS cells, CHO cells, HEK293 cells), insect cells (*e.g.*, *Drosophila melanogaster* S2 cells, *Spodoptera frugiperda* Sf9 cells, High-5 cells), yeast cells, bacterial cells (e.g., *E. coli*) and helminthic cells. The suitability of a particular cell for use as a host cell in accordance with the invention will depend on the ability to introduce a polynucleotide of the invention into the cell by any known means of transformation or transfection. Preferred host cells will also be capable of stably maintaining the introduced polynucleotide and will present a minimum of obstacles to propagation.

Another aspect of the invention is directed to a genetically modified invertebrate comprising a polynucleotide encoding a heterologous GPCR-like receptor (e.g., transgene) as described above or comprising a modification in a native gene encoding a GPCR-like receptor (*e.g.*, an insertional disruption or deletion of the gene). The GPCR-like receptor encoding gene may be expressed at normal levels for that gene or may be overexpressed or underexpressed. A preferred invertebrate for generation of genetically modified organisms is a member of the helminths.

Yet another aspect of the invention is drawn to diagnostic methods for determining neurological abnormalities associated with aberrant GPCR-like receptor activity. Such methods specifically measure the presence, and optionally quantity, of a GPCR-like receptor polynucleotide according to the invention or the presence, and optionally quantity or activity, of a GPCR-like receptor. Any method known in the art may be used to measure the specific polynucleotides or polypeptides of the invention, and measurements may be performed on intact organisms, isolated tissues, or cell cultures. In a related aspect, the invention contemplates methods for diagnosing invertebrate infestation of an organism (*e.g.*, mammals such as humans, mammalian livestock, other vertebrates such as fish, and non-pest invertebrates such as molluscs) or

an environment using a specific measurement of a polynucleotide or polypeptide according to the invention.

Another aspect of the invention is directed to treatment methods. The invention contemplates methods of killing or inhibiting the viability of invertebrates using the modulators described above or identified as described above, including antibodies and antisense polynucleotides. Such methods include methods for treating infestations and/or infections caused by a variety of invertebrate pests, including both ecto- and endo-parasites. A variety of human and other animal ailments, particularly those ailments relating to aberrant neurological functioning, are treated by administering a biologically effective amount of a modulator, GPCR-like receptor polynucleotide or GPCR-like receptor to a cell, tissue, organism, or environment using techniques known in the art.

Use of such modulators in the preparation of a medicament for treating parasitic infection is also contemplated. A variety of administration regimens known in the art are available to deliver a therapeutically effective (*i.e.*, a level of activity capable of deleteriously affecting at least one biological process of a parasite, or more generally, of a pest organism) activity level of a modulator, directly or indirectly, to a parasite or pest, or to an environment associated with the parasite or pest.

Numerous additional aspects and advantages of the invention will become apparent to those skilled in the art upon consideration of the following detailed description of the preferred embodiments of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention has several aspects, one of which is modulating the activity of invertebrate G protein-coupled receptors, which is useful in controlling a wide variety of pest organisms that affect the health of humans and other animals such as domestic pets and livestock, as well as plants. The G protein-coupled receptors typically transduce signals involved in the neurophysiological functioning of the target invertebrates and interference with that functioning frequently proves fatal to the target organism.

Moreover, these receptors are found only in invertebrates, so that specific modulators are safe for use around humans, as well as their pets, livestock, and crops.

A "GPCR-like receptor" is a polypeptide receptor exhibiting the structural characteristics of G protein-coupled receptors (GPCRs), *i.e.*, an N-terminal extracellular region, comprising several loop-like domains (typically three), a transmembrane region comprising seven transmembrane domains arranged in a typical serpentine disposition, and an intracellular region comprising several loop-like domains (typically three).

Invertebrate neuropeptides show profound neuromuscular effects and FaRPs represent the largest family of such neuropeptides known to date. A "FaRP" is a "FMRFamide-related peptide." A "FMRFamide," in turn, is a relatively small peptide, typically having 4-15 amino acids, that matches at least three of the four listed amino acids (FMRF) at its C-terminus. Generally, FaRPs exhibit neurophysiological effects and are therefore properly grouped in a class of neuropeptides that frequently function as ligands of GPCR-like receptors. A "modulator" is an effector of a GPCR-like receptor function; which include the binding of one or more ligands, localization to a membrane, and signal transduction. Signal transduction may involve a change in the relative affinities of a GPCR-associated G protein for various guanidylate nucleotides. With respect to polynucleotides and polypeptides of the invention, "synthesized," as used herein and understood in the art, refers to polynucleotides or polypeptides produced by purely chemical, as opposed to enzymatic, methods. "Wholly" synthesized sequences are therefore produced entirely by chemical means, and "partially" synthesized sequences embrace those wherein only portions of the resulting polynucleotide or polypeptide were produced by chemical means.

The present invention provides purified and isolated polynucleotides (*e.g.*, DNA sequences and RNA transcripts, both sense and complementary antisense strands, including splice variants thereof) encoding invertebrate G protein-coupled receptors. DNA polynucleotides of the invention include genomic DNA, cDNA, and DNA that has been chemically synthesized in whole or in part.

Genomic DNA of the invention comprises the protein coding region for a polypeptide of the invention and includes allelic variants thereof. It is widely understood that, for many genes, genomic DNA is transcribed into RNA transcripts that undergo one

or more splicing events wherein intron (*i.e.*, non-coding regions) of the transcripts are removed, or "spliced out." RNA transcripts that can be spliced by alternative mechanisms, and therefore be subject to removal of different RNA sequences but still encode the same polypeptide, are referred to in the art as splice variants. Splice variants therefore are encoded by the same original genomic DNA sequences but arise from distinct mRNA transcripts found in at least one cell. By way of non-limiting example, several embodiments of the invention are characterized by one of the particular splice variants identified in Table 1, below.

Table 1

SEQ ID NO	Name	Wormpep sequence	Difference(s)
1	CEGPCR1a[1]	AC7.1	94 bp deletion between nucleotides 594-595 of SEQ ID NO:1
3	CEGPCR1f	AC7.1	94 bp deletion between nucleotides 594-595 of SEQ ID NO:3; addition of nucleotides 868-912 of SEQ ID NO:3
5	CEGPCR12c	F41E7.3	90 bp addition at 5' end of SEQ ID NO:5; deletion of 52 bp between nucleotides 173-174 of SEQ ID NO:5; deletion of 6 bp between nucleotides 1115-1116 of SEQ ID NO:5
7	CEGPCR12h	F41E7.3	90 bp addition at 5' end of SEQ ID NO:7; deletion of 52 bp between nucleotides 173-174 of SEQ ID NO:7; deletion of 12 bp between nucleotides 1115-1116 of SEQ ID NO:7
9	CEGPCR12u	F41E7.3	90 bp addition at 5' end of SEQ ID NO:9; deletion of 52 bp between nucleotides 173-174 of SEQ ID NO:9
11	CEGPCR12v	F41E7.3	90 bp addition at 5' end of SEQ ID NO:11; deletion of 52 bp between nucleotides 173-174 of SEQ ID NO:11; addition of 97 bp between nucleotides 1115-1212 of SEQ ID NO:11

[1] The "CEGPCR" labeling system used herein corresponds to the "PNU" labeling system used in related USSN 60/162,523 (*e.g.*, CEGPCR1a corresponds to PNU1a).

SEQ ID NO	Name	Wormpep sequence	Difference(s)
17	CEGPCR18a	Y54E2A.1	deletion of 58 bp between nucleotides 1089-1090 of SEQ ID NO:17; deletion of 32 bp between nucleotides 1331- 1332 of SEQ ID NO:17
106	CEGPCR19.1	Y58G8a.1	
104	CEGPCR19.2	Y58G8a.2	
114	CEGPCR24a	Y59H11A 1.a-01	
116	CEGPCR24b	Y59H11A 1.a-02	

Allelic variants are modified forms of a wild-type gene sequence, the modification resulting from recombination during chromosomal segregation or exposure to conditions which give rise to genetic mutation. Allelic variants, like wild type genes, are naturally occurring sequences (as opposed to non-naturally occurring variants which arise from *in vitro* manipulation), and are also comprehended by the invention.

The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding invertebrate GPCR-like receptors (conventionally followed by second strand synthesis of a complementary strand to provide a double-stranded DNA). In addition to cDNA forms of the polynucleotides identified in Table 1 as splice variants, preferred cDNAs according to the invention include those polynucleotides having sequences selected from the group consisting of SEQ ID NOS: 13, 15 and 21. These sequences differ from sequences in the Wormpep database that exhibit some similarity thereto, as shown in Table 2.

Table 2

SEQ ID NO	Name	Wormpep sequence	Difference(s)
	CEGPCR1a		

SEQ ID NO	Name	Wormpep sequence	Difference(s)
	cegpcr1(a,f), cegpcr4, cegpcr5, cegpcr11, cegpcr12(c,h, u,v), cegpcr13, cegpcr14, cegpcr18, cegpcr19(L,S) , cegpcr22, cegpcr24(a,b)		
21	CEGPCR4	C16D6.2	deletion of 55 bp between nucleotides 1070-1071 of SEQ ID NO:21
	CEGPCR5		
	CEGPCR11		
	CEGPCR12c		
	CEGPCR12h		
	CEGPCR12u		
	CEGPCR12v		
13	CEGPCR13	T05A1.1	deletion of 85 bp between nucleotides 753-754 of SEQ ID NO:13; addition of 226 bp between nucleotides 926-1152 of SEQ ID NO:13
15	CEGPCR14	ZC412.1	deletion of 92 bp between nucleotides 1277-1278 of SEQ ID NO:15
17	CEGPCR18a	Y54E2A.1 a	
106	CEGPCR19.1	Y58G8a.1	
104	CEGPCR19.2	Y58G8a.2	
110	CEGPCR22	C06G4.5	
114	CEGPCR24a	Y59H11A L.a	

SEQ ID NO	Name	Wormpep sequence	Difference(s)
116	CEGPCR24b	Y59H11A L.b	

Preferred DNA sequences encoding invertebrate GPCR-like receptor polypeptides are set out in SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176 and 178. A preferred DNA of the invention comprises a double stranded molecule (for example, the molecule having one of the sequences set forth in the above-referenced SEQ ID NOs, along with the complementary molecule (the "non-coding strand" or "complement") having a sequence unambiguously deducible from one of those sequences, according to Watson-Crick base-pairing rules for DNA). Also preferred are other polynucleotides encoding one of the invertebrate GPCR-like receptor polypeptides of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 105, 107, 109, 111, 113, 115, 117, 177, and 179 which may differ in sequence from the corresponding polynucleotides of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176 and 178 by virtue of the well-known degeneracy of the universal genetic code.

The invention further embraces invertebrate species (preferably helminth and insect) homologs of the disclosed GPCR-like DNAs. Species homologs, sometimes referred to as "orthologs," in general share at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% similarity with the DNA sequences disclosed herein. Percent sequence "similarity" with respect to polynucleotides of the invention is defined herein as the percentage of nucleotide bases in the candidate sequence that are identical to nucleotides in the relevant sequences set forth in the sequence listing below, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Preferred for comparative sequence analyses is the BLAST

program available from GCG, implemented with default parameters. Also preferred is the Blastall program available from NCBI.

The polynucleotide sequence information provided by the invention makes possible large scale expression of an encoded polypeptide by techniques well known and routinely practiced in the art. Polynucleotides of the invention also permit identification and isolation of polynucleotides encoding related GPCR-like polypeptides, such as the aforementioned allelic variants and species homologs, by well known techniques including Southern and/or Northern hybridization, and by polymerase chain reaction (PCR). Examples of related polynucleotides include polynucleotides encoding polypeptides homologous to the invertebrate GPCRs and structurally related polypeptides sharing one or more biological, immunological, and/or physical properties of those GPCRs. Invertebrate genes encoding proteins homologous to the disclosed GPCRs can also be identified by Southern and/or PCR analysis and are useful in methods of the invention described below. Knowledge of the sequence of the disclosed GPCR-encoding DNAs also makes possible, through use of Southern hybridization, polymerase chain reaction (PCR), and other known techniques, the identification of genomic DNA sequences encoding GPCR expression control regulatory sequences such as promoters, operators, enhancers, repressors, and other regulatory sequences known in the art. Polynucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express an invertebrate GPCR.

The disclosure herein of a full-length polynucleotide encoding invertebrate GPCR-like polypeptides makes readily available to the worker of ordinary skill in the art every possible fragment of the full-length polynucleotides. The invention therefore provides fragments of invertebrate GPCR-encoding polynucleotides comprising at least 14-15, and preferably at least 18, 20, 25, 50, or 75 consecutive nucleotides of a polynucleotide encoding an invertebrate GPCR.

Fragment polynucleotides contemplated by the invention encode immunologically active peptides capable of interacting with specific anti-GPCR antibodies, as well as domains of GPCRs. The GPCR domains (*i.e.*, the N-terminal extracellular domain, one or more of the three extracellular loop domains, one or more of the seven transmembrane domains, one or more of the three intracellular loop domains, or the C-

terminal intracellular domain) of invertebrate GPCRs disclosed herein are characterized in Example 2 and Table 6, below. Full length or fragment polynucleotides can be linked to polynucleotides encoding heterologous polypeptides, *e.g.*, for producing variants that are fusion proteins.

Preferably, fragment polynucleotides of the invention comprise sequences unique to one of the disclosed GPCR-encoding sequences, and therefore hybridize under highly stringent or moderately stringent conditions only (*i.e.*, "specifically") to polynucleotides encoding a disclosed invertebrate GPCR (or fragments thereof) and not to polynucleotides encoding other GPCRs. Polynucleotide fragments of genomic sequences of the invention comprise not only sequences unique to the coding region, but also fragments of the full-length sequence derived from introns, regulatory regions, and/or other non-translated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, *e.g.*, those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic DNA to which a polynucleotide will hybridize. Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent, and enzymatic labeling.

Fragment polynucleotides are particularly useful as probes for detection of full-length invertebrate GPCR-like polynucleotides, or other polynucleotide fragments thereof. One or more fragment polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding an invertebrate GPCR-like receptor polypeptide, or used to detect variations in a polynucleotide sequence encoding such a polypeptide.

The invention also embraces DNAs encoding invertebrate GPCR-like polypeptides that hybridize under conditions of moderate or high stringency to the non-coding strand, or complement, of a polynucleotide comprising one of the sequences set forth in SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176 and 178.

Exemplary conditions of high stringency are as follows: hybridization at 42°C in a solution (*i.e.*, a hybridization solution) comprising 50% formamide, 1% SDS, 1

M NaCl, 10% Dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS. It is understood in the art that conditions of equivalent stringency can be achieved through variation of temperature and buffer, or salt concentration, as described in Ausubel, *et al.* (Eds.), Protocols in Molecular Biology, John Wiley & Sons (1994), pp. 6.0.3 to 6.4.10. Modifications in hybridization conditions can be empirically determined or precisely calculated based on the length and the percentage of guanosine/cytosine (GC) base pairing of the probe. The hybridization conditions can be calculated as described in Sambrook, *et al.*, (Eds.), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47 to 9.51.

The invention also provides a purified and isolated invertebrate GPCR-like polypeptide encoded by a polynucleotide of the invention. Presently preferred is a polypeptide comprising one of the amino acid sequences set out in SEQ ID NOs : 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24,26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 105, 107, 109, 111, 113, 115, 117, 177, and 179, and species homologs thereof. The invention also embraces a GPCR-like polypeptide encoded by a DNA selected from the group consisting of : (a) the DNA sequence set out in any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25,27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 104, 106, 108, 110, 112, 114, 116, 176 and 178 and species homologs thereof; and (b) a DNA molecule encoding a GPCR-like gene product that hybridizes under conditions of moderate or high stringency to the DNA of (a). The invention further embraces polypeptides that have at least about 99%,at least about 95%, at least about 90%, at least about 85%, at least about 80%, at least about 75%, at least about 70%, at least about 65%, at least about 60%, at least about 55%, and at least about 50% identity and/or homology to the preferred polypeptides of the invention. Percent amino acid sequence "identity" with respect to the preferred polypeptides of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the GPCR-like gene product sequence after aligning both sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Percent sequence "homology" with respect to the preferred polypeptides of the invention is defined herein as

the percentage of amino acid residues in the candidate sequence that are identical with the residues in one of the GPCR-like receptor polypeptide sequences after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and also considering any conservative substitutions as part of the sequence identity. Conservative substitutions are defined as set out in Tables 3 and 4.

GPCR polypeptides of the invention may be isolated from natural invertebrate cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention. GPCR-like gene products of the invention may be full-length polypeptides, biologically active fragments, or variants thereof which retain specific biological or immunological activity. Variants may comprise GPCR-like polypeptide analogs wherein one or more of the specified (*i.e.*, naturally encoded) amino acids is deleted or replaced or wherein one or more non-specified amino acids are added: (1) without loss of one or more of the biological activities or immunological characteristics specific for the GPCR-like receptor; or (2) with specific disablement of a particular biological activity of the GPCR-like polypeptide. Contemplated deletion variants also include fragments lacking portions of a GPCR-like polypeptide not essential for biological activity, and insertion variants include fusion polypeptides in which the wild-type GPCR-like polypeptide or fragment thereof has been fused to another polypeptide.

Table 3

Conservative Substitutions I

SIDE CHAIN CHARACTERISTIC		AMINO ACID
Aliphatic	Non-polar	G A P I L V
	Polar - uncharged	C S T M N Q
	Polar - charged	D E K R
Aromatic		H F W Y
Other		N Q D E

Table 4

Conservative Substitutions II

SIDE CHAIN CHARACTERISTIC	AMINO ACID
Non-polar (hydrophobic)	
A. Aliphatic:	A L I V P
B. Aromatic:	F W Y
C. Sulfur-containing:	M
D. Borderline:	G
Uncharged-polar	
A. Hydroxyl:	S T Y
B. Amides:	N Q
C. Sulfhydryl:	C
D. Borderline:	G
Positively Charged (Basic):	K R H
Negatively Charged (Acidic):	D E

Variant GPCR-like polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Conservative substitutions are recognized in the art to classify amino acids according to their related physical properties and can be defined as set out in Table 3. (See, WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96).) Alternatively, conservative amino acids can be grouped as defined in Lehninger, [Biochemistry, Second Edition; Worth Publishers, Inc. NY:NY (1975), pp.71-77] as set out in Table 4.

Variant GPCR-like polypeptides of the invention include mature GPCR-like gene products, *i.e.*, wherein leader or signal sequences are removed, having additional amino terminal residues. GPCR-like gene products having an additional methionine residue at position -1 are contemplated, as are GPCR-like receptors having additional methionine and lysine residues at positions -2 and -1. Variants of these types are particularly useful for recombinant protein production in bacterial cell types. Variants of the invention also include gene products wherein amino terminal sequences derived from other proteins have been

introduced, as well as variants comprising amino terminal sequences that are not found in naturally occurring proteins.

Another component useful in some methods for identifying modulators of GPCR-like receptor activity, described below, is a class of neuropeptides that includes FaRPs. The FaRPs comprise a large family of peptides that typically function as neurotransmitters in the invertebrate nervous system. The FaRPs that are associated with these neurophysiological functions generally are about four to about nine amino acid residues in length and include a C-terminal amino acid sequence motif of (aromatic)-X-R-Famide, where X is typically M, L, I or V. Day *et al.*, Peptides 20:999-1019 (1999), which is incorporated by reference in its entirety. However, peptides that are exceptions to this rule are also contemplated by the invention, particularly those that are co-encoded on FaRP precursor genes or those that share the characteristics of conforming peptides. Also contemplated by the invention are neuropeptide variants, including FaRP variants, peptides exhibiting an RYamide motif, and other peptides such as retro-inverso neuropeptides (*e.g.*, FaRPs comprising the D stereoisomers of amino acids in a sequence that is reversed from a reference FaRP). For example a retro-inverso variant of A_L-F_L-M_L-R_L-F_L would be F_D-R_D-M_D-F_D-A_D.

Neuropeptides such as FaRPs have been found in a wide variety of invertebrates, including arthropods such as insects (*e.g.*, locusts and flies such as *Drosophila*) and lobsters, mollusks such as the snail *Lymnaea stagnalis*, and helminths such as nematodes (*e.g.*, *C. elegans*, *A. suum*, *Haemonchus contortus*), trematodes (*e.g.*, *Schistosoma mansoni*) and cestodes, among others, including Manduca. Information relating to the structure, function, and structure-function relationships of these neuropeptides is known in the art [Day *et al.* (1999)]. Table 5 below provides a classification of several known *C. elegans* FaRPs, which are examples of the class of neuropeptides useful in practicing the invention.

Table 5

***C. elegans* FMRFamide-related peptides (FaRPs) and encoding genes**

GENE AND CHROMOSOME NUMBER	STRUCTURALLY CHARACTERIZED FaRPs	PEPTIDES PREDICTED FROM GENE
FLP-1	SDPNFLRFa SADPNFLRFa SQPNFLRFa ASGDPNFLRFa AAADPNFLRFa	KPNFMRYa AGSDPNFLRFa SQPNFLRFa ASGDPNFLRFa SDPNFLRFa AAADPNFLRFa SADPNFLRFa KPNFLRFa
FLP-2		LRGEPIRFa SPREPIRFa
FLP-3		SPLGTMRFa TPLGTMRFa SAEPFGTMRFa NPENDTPFGTMRFa ASEDALFGTMRFa EDGNAPFGTMRFa EAEEPLGTMRFa SADDSAPFGTMRFa NPLGTMRFa
FLP-4		PTFIRFa ASPSFIRFa
FLP-5		APKPKFIRFa AGAKFIRFa GAKFIRFa
FLP-6	KSAYMRFa	(6 copies) KSAYMRFa
FLP-7		(2 copies) TPMQRSSMVRFa (3 copies) SPMQRSSMVRFa SPMERSAMVRFa SPMDRSKMVRFa
FLP-8		(3 copies) KNEFIRFa
FLP-9	KPSFVRFa	(2 copies) KPSFVRFa
FLP-10		QPKARSGYIRFa

FLP-11		AMRNALVRFa ASGGMRNALVRFa NGAPQPFVRFa
FLP-12		RNKFEFIRFa
FLP-13		SDRPTRAMDSPLIRFa (2 copies) AADGAPLIRFa (2 copies) APEASPFIRFa ASPSAPLIRFa SPSAVPLIRFa SAAAPLIRFa ASSAPLIRFa
FLP-14		(4 copies) KHEYLRFa
FLP-15		GGPQGPLRFa GPSGPLRFa
FLP-16		AQTFVRFa GQTFVRFa
FLP-17		(2 copies) KSAFVRFa KSQYIRFa
FLP-18		DFDGAMPGVLRFa DMPGVLRFa KSVPGVLRFa SVPGVLRFa EIPGVLRFa SEVPGVLRFa DVPGVLRFa SVPGVLRFa
OTHER PUTATIVE FLP GENES		TKFQDFLRFa AMRNSLVRFa DYDFVRFa DGFVRFa AFFKNVLRFa

"a" means amide.

Further characterization of the structure-function relationships of neuropeptides, such as FaRPs, that are embraced by the invention may be readily accomplished by one of ordinary skill in the art. For example, amino acid-scan modifications (*e.g.*, alanine-scan), in which each residue is sequentially replaced with

another amino acid such as alanine are available. Additionally, given the knowledge in the art (see generally, Geary et al., 1999), such modifications to known neuropeptides (*e.g.*, FaRPs) as the internal or terminal deletion of one or more amino acids, as well as the internal or terminal addition of residues, involves no more than routine experimentation. Such "variants" of FaRPs and related neuropeptides are among the neuropeptides contemplated for use in the methods of the invention.

The GPCR-like receptor polypeptides and neuropeptides produced by the methods described above are useful in assays for modulators of GPCR-like receptor activities, including binding partner (*e.g.*, ligand) binding. Assays contemplated by the invention include both binding assays and activity assays; these assays may be performed in conventional or high throughput formats. GPCR-like receptor activity is defined as including the binding of any binding partner, such as a ligand, as well as the propagation of any transmembrane signal (*e.g.*, stimulation of a G protein or influence on the flux of an ion across a membrane). Modulator screens are designed to identify stimulatory and/or inhibitory agents. The sources for potential agents to be screened include natural sources, such as a cell extract (*e.g.*, invertebrate cells including, but not limited to, bacterial, fungal, algal, and plant cells) and synthetic sources, such as chemical compound libraries. For proteins with known activity, function assays are established based on the activity, and a large number of potential agents are screened for the ability to either stimulate or inhibit the activity. Binding assays are used to detect GPCR-like receptor binding activity to neuropeptide or non-peptidic ligands. Both functional and binding assays of GPCR-like receptor activity are readily adapted to screens for modulators such as inhibitory compounds.

The invention contemplates a multitude of assays to screen and identify inhibitors of ligand binding to GPCR-like receptors. In one example, the GPCR-like receptor is immobilized and interaction with a binding partner is assessed in the presence and absence of a candidate modulator such as an inhibitor compound. In another example, interaction between the GPCR-like receptor and its binding partner is assessed in a solution assay, both in the presence and absence of a candidate inhibitor compound. In either assay, an inhibitor is identified as a compound that decreases binding between

the GPCR-like receptor and its binding partner. Another contemplated assay involves a variation of the di-hybrid assay wherein an inhibitor of protein/protein interactions is identified by detection of a positive signal in a transformed or transfected host cell.

Candidate modulators contemplated by the invention include any chemical compounds, including libraries of chemical compounds. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, or analogs of known compounds, or analogs of compounds that have been identified as "hits" or "leads" in prior drug discovery screens, some of which may be derived from natural products or from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms which are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see Science 282:63-68 (1998). Combinatorial libraries are composed of large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or synthetic methods. Of particular interest are non-peptide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol. 8*:701-707 (1997). Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate activity.

Candidate modulators contemplated by the invention can be designed and include soluble forms of binding partners, as well as chimeric, or fusion, proteins thereof. A "binding partner" as used herein broadly encompasses non-peptide modulators, peptide modulators (e.g., neuropeptide variants), antibodies (including monoclonal and

polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR and/or antigen-binding sequences, which specifically recognize a polypeptide of the invention), antibody fragments; and modified compounds comprising antibody domains that are immunospecific for the expression product of the identified GPCR-like gene.

A number of assays are known in the art that can identify chemical compounds that bind to or interact with a GPCR-like receptor. Such assays are useful, for example, in methods of identifying candidate modulators described herein, or in methods for identifying specific neuropeptide ligands of a GPCR-like receptor. Assays that measure binding or interaction of compounds with target proteins include assays that identify compounds that inhibit unfolding or denaturation of a target protein, assays that separate compounds that bind to target proteins through affinity ultrafiltration followed by ion spray mass spectroscopy/HPLC methods or other physical and analytical methods, capillary electrophoresis assays and two-hybrid assays.

One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (*i.e.*, when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of the ligand to the target protein can be determined by any method which distinguishes between the folded and unfolded states of the target protein. The function of the target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

Another method for identifying ligands of a target protein is described in Wieboldt *et al.*, Anal. Chem., 69:1683-1691 (1997), incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields *et al.*, Nature, 340:245-246 (1989), and Fields *et al.*, Trends in Genetics, 10:286-292 (1994), both of which are incorporated herein by reference. The two-hybrid system is a genetic assay for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain, cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene.

When the function of the GPCR-like receptor is unknown and no ligands are known to bind the gene product, the yeast two-hybrid assay can be used to identify proteins that bind to the receptor. In an assay to identify proteins that bind to a GPCR-like receptor, or fragment thereof, a fusion polynucleotide encoding both a GPCR-like receptor or fragment (*i.e.*, a first protein) and a UAS binding domain may be used. In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

In addition, when the GPCR-like receptor or fragment thereof is known to interact with another protein or nucleic acid, the two-hybrid assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent, for example, results in lack of or reduction in a reporter signal.

The candidate modulators identified by the initial screens are evaluated for their effect on neuromuscular function using *in vivo* or *ex vivo* systems. A preferred evaluation method exposes the candidate modulators to *A. suum* neuromuscular strips and records neuromuscular activity relative to control strips exposed to compounds other than the candidate modulator under investigation. Other methods of screening for modulators are described in the following examples, and still other screening methods are well known in the art. Other aspects and advantages of the present invention will be understood upon consideration of the following illustrative examples. Example 1 describes the cloning of GPCR-like receptors; Example 2 provides an analysis of the polynucleotide sequences encoding those receptors in the context of their structural features such as regions and domains; Example 3 discloses GPCR-like receptor

expression studies; Example 4 describes antibodies to GPCR-like receptors; Example 5 provides a variety of assays to identify modulators of GPCR-like receptor activity; and Example 6 describes approaches to the identification of candidate GPCR-like receptors, as well as the cloning and characterization of such receptors.

Example 1

Cloning of GPCR-Like Receptors

A. Database search

The Wormpep database, containing all of the predicted protein sequences encoded by the *C. elegans* genome, used in these studies, was obtained through the Sanger Centre Web site (http://www.sanger.ac.uk/Projects/C_elegans/). Wormpep versions 13 (2/13/98) through 23 (released September 4, 2000). This database contains 19,430 protein sequences, including 388 splice variants. *C. elegans* genomic DNA sequences were accessed through ACEDB (Release WS3 4-25). The databases were searched and manipulated using programs from the Wisconsin Package GCG programs.

B. Subcloning of the Coding Region via PCR

Standard molecular biology techniques were performed as described in Ausubel, *et al.* (Eds.), Protocols in Molecular Biology, John Wiley & Sons (1994). Based on the database analyses described above, 22 candidate GPCR-like receptor genes were chosen for PCR amplification. The sequences of primers used for PCR amplification are provided in SEQ ID NOS: 47-76. PCR primers were designed to incorporate an optimized translation initiation sequence (Kozak *et al.*, Nucl. Acids Res. 15(20):8125-8148 (1987)) around the initiator methionine codon and to incorporate a unique restriction site at the 3' end for subsequent linearization of the PCR-amplified fragments. The Expand™ High Fidelity PCR System (Boehringer-Mannheim Corp., Indianapolis, IN), including a mixture of Taq and Pwo DNA polymerases, was used for the PCR experiments. Reaction volumes of 100 µl contained 2.6 Units of enzyme mix ("Units"defined by the supplier), 200 µM of each standard dNTP, 1X Expand™ HF buffer with 15 mM $MgCl_2$, 300 nM of each primer, and 2 µl (approximately 10 ng) of a

DNA template preparation, described below. DNA template preparations were the products of a first-strand cDNA synthesis reaction (heat-inactivated after completion) catalyzed by SuperScript reverse transcriptase (GIBCO-BRL, Gaithersburg, MD) with random hexamers to prime the synthesis using a preparation of total *C. elegans* RNA as template. PCR amplifications were performed in a Perkin-Elmer 9600 thermocycler using the following conditions: 2 minutes at 94°C; 30 cycles of 15 seconds at 94°C, 30 seconds at 50°C, 2 minutes at 70°C; and a final incubation at 70°C for 7 minutes. PCR products were electrophoresed through 0.8% agarose/TAE gels and visualized using ethidium bromide.

Amplified PCR products were cloned using the pCR2.1 TA-tailed vectors (Stratagene, La Jolla, CA). PCR products were excised from the agarose gels and purified using QIAquick columns (Qiagen, Valencia, CA). Ligation reactions were performed in accordance with the recommendations of the supplier (Stratagene). Ligation reactions were transformed into *E. coli* DH5α competent cells and plated on LB plates containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). White colonies were picked and examined for the presence of the appropriate size insert by a commercial mini-prep DNA isolation procedure (Qiagen). To facilitate subsequent expression studies, all cloned receptors were subcloned into the mammalian expression vector, pCR3.1 (Stratagene). Two microgram aliquots of vector and each GPCR-like receptor clone were digested using either *Bst*XI (clones CEGPCR1a, CEGPCR1f; each showing some sequence similarity to Wormpep AC7.1) or *Eco*RI (clones CEGPCR3 (Wormpep C10C6.2), CEGPCR7 (Wormpep C39E6.6), CEGPCR8 (Wormpep C50F7.1), CEGPCR9 (Wormpep C56G3.1), CEGPCR12h and CEGPCR12u (each having some sequence similarity to Wormpep F41E7.3), CEGPCR13 (having some sequence similarity to Wormpep T05A1.1), and CEGPCR15 (Wormpep C49A9.7)). Restriction endonucleases were purchased from New England Biolabs (Beverly, MA) and were used according to the manufacturer's instructions. Following digestion, the linear vector and the GPCR-like receptor sequences were isolated from 1% TAE agarose gels and purified using a QIAquick gel extraction kit (Qiagen). Ligation reactions were performed at 14°C for 16 hours and contained 100 ng vector DNA, 100 ng GPCR-like

insert DNA and 1 Unit T4 DNA ligase (Boehringer Mannheim). Plasmids were propagated in *E. coli* strain DH5α, then isolated and purified using a Qiagen column. Subclones were identified and insert orientation determined by restriction endonuclease analysis as follows: CEGPCR1a and CEGPCR15 (SEQ ID NOS: 1 and 33), *Hin*dIII; CEGPCR3 and CEGPCR8, (SEQ ID NOS: 43 and 27) *Ava*I; CEGPCR7, (SEQ ID NO:25) *Bam*HI; CEGPCR13, (SEQ ID NO:13) *Hin*cII/*Hin*dIII; CEGPCR1f, (SEQ ID NO:3) *Bam*HI/*Bgl*II; and CEGPCR9, CEGPCR12h and CEGPCR12u, (SEQ ID NOS: 29, 7 and 9) *Xba*I.

Example 2

Analysis of the GPCR-Like Receptor Sequences

DNA from pCR2.1 recombinant plasmids containing the expected insert size were sequenced using BigDye dye terminator chemistry (PE Applied Biosystems, Foster City, CA) on an ABI 377 automated DNA sequencer. Complete sequences were assembled using Sequencher software (GeneCodes, Ann Arbor, MI) and compared against the expected sequence as predicted from Wormpep. Each of the 22 putative GPCR-like receptor clones appeared to contain a full-length coding region, although some of the encoded GPCR-like receptor sequences could not be confirmed as having a sequence reported in the Wormpep database. In particular, clones CEGPCR1a, CEGPCR1f, CEGPCR4, CEGPCR12c, CEGPCR12h, CEGPCR12u, CEGPCR12v, CEGPCR13, CEGPCR14, and CEGPCR18a are known to differ from sequences found in the Wormpep database (see Tables 1 and 2, above), while clones CEGPCR5, CEGPCR6, CEGPCR10, CEGPCR11, CEGPCR19 and CEGPCR20 are not known to have sequences found in the Wormpep database.

All 22 of the cloned GPCR-like receptor sequences exhibited the structural characteristics of GPCRs. Each coding region encoded seven hydrophobic stretches of amino acids consistent with transmembrane domains, the characteristic N-terminal extracellular domain, C-terminal cytoplasmic domain and both extracellular and intracellular loops. Table 6 lists the endpoints of all seven transmembrane domains in each of the 22 GPCR-like receptors. In identifying the transmembrane domains,

moreover, the N-terminal extracellular domains are effectively identified (the start of the amino acid sequence to the start of the first transmembrane domain), the three extracellular loop domains (the end of 2TM to the start of 3TM, the end of 4TM to the start of 5TM, and the end of 6TM to the start of 7TM, respectively, in Table 6), the three intracellular loop domains (the end of 1TM to the start of 2TM, the end of 3TM to the start of 4TM, and the end of 5TM to the start of 6TM, respectively, in Table 6), and the C-terminal intracellular domain (the end of 7TM to the end of the amino acid sequence).

Table 6

SIN	TM1		TM2		TM3		TM4		TM5		TM6		TM7	
	start	end	start	end	start	end	start	end	start	end	start	end	start	end
CEGPCR1a	47	71	83	109	121	142	161	181	213	233	262	282	298	321
CEGPCR1f	47	71	83	109	121	142	161	181	213	233	277	297	313	336
CEGPCR2	42	66	78	104	116	137	158	178	204	224	320	340	351	374
CEGPCR3	26	50	62	88	100	121	140	160	196	216	252	272	285	308
CEGPCR4	28	52	64	90	102	123	142	162	192	212	256	276	293	316
CEGPCR5	40	64	76	102	114	135	154	174	212	232	272	292	305	328
CEGPCR6	12	36	48	74	85	106	126	146	178	198	229	249	263	286
CEGPCR7	27	51	63	89	101	122	141	161	192	212	278	298	322	345
CEGPCR8	30	54	66	92	104	125	146	166	194	214	244	264	283	306
CEGPCR9	51	75	87	113	125	146	167	187	239	259	330	350	368	391
CEGPCR11	25	49	61	87	99	120	139	159	189	209	246	266	286	309
CEGPCR12c	38	62	74	100	112	133	157	177	207	227	255	275	291	314
CEGPCR12h	38	62	74	100	112	133	157	177	207	227	255	275	291	314
CEGPCR12u	38	62	74	100	112	133	157	177	207	227	255	275	291	314
CEGPCR12v	38	62	74	100	112	133	157	177	207	227	255	275	291	314
CEGPCR13	27	51	63	89	101	122	141	161	190	210	276	296	320	343
CEGPCR14	30	54	65	91	103	124	143	163	198	218	254	274	299	322
CEGPCR15	62	86	98	124	136	157	176	196	235	255	285	305	321	344
CEGPCR16	29	53	65	91	103	124	143	163	193	213	250	270	287	310
CEGPCR17	36	60	72	98	110	131	150	170	220	240	274	294	306	329
CEGPCR18a	10	34	46	72	84	105	126	146	185	205	282	302	320	343
CEGPCR19.1	56	80	92	118	130	151	170	190	228	248	287	307	324	347
CEGPCR19.2	56	80	92	118	130	151	170	190	228	248	287	307	324	347
CEGPCR20	44	68	79	105	117	138	159	179	215	235	283	303	323	346
CEGPCR21	76	100	112	138	151	172	193	213	249	269	336	356	373	396
CEGPCR22	55	79	95	121	132	153	173	193	233	253	279	299	319	342
CEGPCR23	21	45	64	90	101	122	142	162	203	223	249	269	287	310
CEGPCR24a	51	75	87	113	128	149	170	189	223	243	279	299	316	339
CEGPCR24b	51	75	87	113	128	149	170	189	223	243	279	299	316	339

Where "SIN" refers to SEQ ID NOS.

Example 3

Recombinant Expression of GPCR-Like Receptors

Recombinant pCR3.1 clones encoding invertebrate GPCR-like receptors were transformed into wild-type CHO-K1 cells and functional expression was achieved using a temperature-shift incubation protocol. The CHO-K1 cells were cultured at 37°C in a humidified atmosphere of 5% CO_2 in air in DMEM medium supplemented with 10% heat-inactivated fetal bovine serum (*i.e.*, FBS), 10 µg/ml gentamicin, and 0.1 mM

nonessential amino acids to give complete DMEM medium. Additionally, for incubations at temperatures lower than 37°C, the medium was fortified with 10 mM HEPES (pH 7.4). Cells were transfected with recombinant pCR3.1 clones containing GPCR-like receptor DNAs, using LipofectAMINE PLUS™, essentially according to the manufacturer's instructions. Briefly, CHO-K1 cells were plated on 10 cm sterile tissue culture dishes (Corning Glass Works, Corning, NY) and were incubated until the plates showed about 50-60% confluent growth. In separate plastic tubes, PLUS (20 µl/plate; Life Technologies) was added to 0.75 ml/plate of recombinant plasmid DNA in OptiMEM (5 µg/plate; Gibco/BRL - Life Technologies), mixed and incubated at room temperature for 15 minutes. Separately, LipofectAMINE (30 µl/plate) was mixed with 0.75 ml OptiMEM and added to the pre-complexed DNA/PLUS mixture and incubated at room temperature for 15 minutes. The culture medium was then replaced with serum-free transfection medium (DMEM, 5 ml/plate). The DNA-PLUS-LipofectAMINE complex was then added to the plates (1.5 ml/plate) and mixed gently into the medium, followed by a 3 hour incubation at 37°C in 5% CO_2. At this time, the medium was supplemented with complete DMEM medium containing 20% FBS (6.5 ml/plate) and incubation was continued at 37°C in 5% CO_2. After a total of 24 hours at 37°C in 5% CO_2, the medium was replaced with complete DMEM medium fortified with 10 mM HEPES (10 ml/plate) and the cells were moved to an incubator maintained at 28-30°C in a humidified atmosphere of 5% CO_2 in air. The incubation continued for an additional 24 hours, typically at the lower temperature indicated above. A plasmid encoding Green Fluorescent Protein (*i.e.*, GFP, 4 µg/plate) was used for transient GFP expression in CHO-K1 cells to estimate the transfection yields under the same conditions as used for GPCR-like receptor polynucleotides.

To assess expression, membranes were prepared from transfected CHO-K1 cells. Transfected cells were washed once with ice-cold Dulbecco's phosphate-buffered saline (*i.e.*, PBS), 5 ml per 10 cm plate, and the cells were scraped into 5 ml of the same buffer. Cell suspensions from multiple plates were combined and centrifuged at 500 x *g* for 10 minutes at 4°C. The cell pellet was reconstituted in ice-cold TEE (25 mM Tris, pH 7.4, 5 mM EGTA, 5 mM EDTA). Convenient aliquots were snap-frozen in liquid nitrogen and stored at -70°C. After thawing, the cells were homogenized and

centrifuged at 500 x *g* for 5 minutes at 4°C to pellet nuclei and unbroken cells. The supernatant was centrifuged at 47,000 x *g* for 30 minutes at 4°C. The membrane pellet was washed once with TEE, resuspended in 20 mM HEPES, pH 7.4, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA (assay buffer), distributed into aliquots of 0.5-1.0 ml, and frozen in liquid nitrogen. Membrane aliquots were stored at -70°C. Membrane protein concentration was determined using the BCA Protein Assay Reagent from Pierce (Rockford, Illinois), with bovine serum albumin (*i.e.*, BSA) as a standard.

Expression of the recombinant GPCR-like receptors was assessed using a function-based [^{35}S]GTPγS assay, which is described in greater detail in Example 5F, below. In general, the post-transfection reduction in incubation temperature led to significant increases in detectable GPCR-like receptor activity. More specifically, a controlled study of three cloned GPCR-like receptors, CEGPCR3 (Wormpep C10C6.2), CEGPCR12c (exhibiting some sequence similarity to Wormpep F41E7.3), and CEGPCR7 (Wormpep C39E6.6), demonstrated that the temperature reduction resulted in significant increases in the expression levels of all three receptors. The GTPγS assay described in Example 5F was also used to screen for neuropeptide (including FaRP) ligands binding to invertebrate GPCR-like receptors. The results are presented in Tables 8-13, below, which identify GPCR-like receptors binding to one or more neuropeptide (*e.g.*, FaRP) ligands, GPCR-like receptors binding to neuropeptide ligands having specified structures, the sources of those particular ligands, and EC_{50} (nM) values. As shown in the tables, such ligands were identified for CEGPCR3, CEGPCR4, CEGPCR5, CEGPCR7, CEGPCR12c, CEGPCR12h, CEGPCR12u, CEGPCR12v, CEGPCR16, CEGPCR19.1, and CEGPCR19.2, and on that basis these receptors have been identified as neuropeptide (FaRP) receptors.

CEGPCR3, CEGPCR4, CEGPCR5, CEGPCR12c and CEGPCR16 were not detectably activated by their matched ligands, GNSFLRFamide (SEQ ID NO:88), PDVDHVFLRFamide (SEQ ID NO:94), GLGPRPLRFamide (SEQ ID NO:86), PDVDHVFLRFamide (SEQ ID NO: 94), and ASEDALFGTMRFamide (SEQ ID NO: 77), respectively, when cell membranes were prepared from cells transfected and maintained at 37°C. In contrast, a temperature shift to 28°C 24 hours after transfection resulted in significant stimulation of ligand-induced [^{35}S]GTPγS binding.

In the case of the CEGPCR7-AF9 system, cell transfections and constant temperature (37°C) incubations were performed three times. No stimulation of the GLGPRPLRFamide (AF9)-induced [^{35}S]GTPγS binding was detected in two of these experiments while in the third one, a 1.4 fold elevated [^{35}S]GTPγS binding was recorded. The extent of this stimulation was, however, considerably smaller than the stimulation in activity achieved by using the cooling (28°C) protocol (2.6 fold elevation in activity). Thus, the temperature downshift was important for the functional expression of *C. elegans* GPCRs expressed in mammalian cells. In contrast, the responses of the human D2 receptor control to its ligand, dopamine, were the same regardless of the transfection protocol used. Moreover, no temperature-dependent receptor activity was detected when the hD2 receptor control was exposed to a number of neuropeptides, including FaRPs, indicating that the temperature-dependent GPCR-like receptor activity was specific to the cloned receptors.

All invertebrate GPCR-like receptors, including those receptors identified in Table 6, are screened against the invertebrate neuropeptides, including many FaRPs. It is expected that each one of the invertebrate GPCR-like receptors will be activated by one or more neuropeptides of the class including FaRPs, using an assay disclosed in Example 5, such as the GTPγS assay.

Example 4

Antibodies to GPCR-Like Receptors

Standard techniques are employed to generate polyclonal or monoclonal antibodies to the GPCR-like receptors, and to generate useful antigen-binding fragments thereof or variants thereof. Such protocols can be found, for example, in Sambrook et al., *Molecular Cloning: a Laboratory Manual.* Second Edition, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory (1989); Harlow *et al.* (Eds), *Antibodies A Laboratory Manual*; Cold Spring Harbor Laboratory; Cold Spring Harbor , NY (1988). In one embodiment, recombinant GPCR-like polypeptides (or cells or cell membranes containing such polypeptides) are used as antigens to generate the antibodies. In another embodiment, one or more peptides having amino acid sequences corresponding to an immunogenic portion of a GPCR-like receptor (*e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

17, 18, 19, 20, or more amino acids) are used as antigen. Peptides corresponding to extracellular portions of GPCR-like receptors, especially hydrophilic extracellular portions, are preferred. The antigen may be mixed with an adjuvant or linked to a hapten to increase antibody production. Polyclonal and monoclonal antibodies, chimeric (*e.g.*, humanized) antibodies, fragments of antibodies, and all other forms of antibody molecules disclosed herein are referred to collectively as antibody products.

A. Polyclonal or Monoclonal Antibodies

As one exemplary protocol, a recombinant GPCR-like polypeptide or a synthetic fragment thereof is used to immunize a mouse for generation of monoclonal antibodies (or larger mammal, such as a rabbit, for polyclonal antibodies). To increase antigenicity, peptides are conjugated to Keyhole Lympet Hemocyanin (Pierce), according to the manufacturer's recommendations. For an initial injection, the antigen is emulsified with Freund's Complete Adjuvant and injected subcutaneously. At intervals of two to three weeks, additional aliquots of GPCR-like receptor antigen are emulsified with Freund's Incomplete Adjuvant and injected subcutaneously. Prior to the final booster injection, a serum sample is taken from the immunized mice and assayed by Western blot to confirm the presence of antibodies that immunoreact with a GPCR-like polypeptide. Serum from the immunized animals may be used as a polyclonal antisera or used to isolate polyclonal antibodies that recognize a GPCR-like receptor. Alternatively, the mice are sacrificed and their spleens are removed for generation of monoclonal antibodies.

To generate monoclonal antibodies, the spleens are placed in 10 ml serum-free RPMI 1640, and single-cell suspensions are formed by grinding the spleens in serum-free RPMI 1640, supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 100 Units/ml penicillin, and 100 µg/ml streptomycin (RPMI) (Gibco, Canada). The cell suspensions are filtered and washed by centrifugation and resuspended in serum-free RPMI. Thymocytes taken from three naive Balb/c mice are prepared in a similar manner and used as a Feeder Layer. NS-1 myeloma cells, kept in log phase in RPMI with 10% (FBS (Hyclone Laboratories, Inc., Logan, Utah) for three days prior to fusion, are centrifuged and washed as well.

To produce hybridoma fusions, spleen cells from the immunized mice are combined with NS-1 cells and centrifuged, and the supernatant is aspirated. The cell pellet is dislodged by tapping the tube, and 2 ml of 37°C PEG 1500 (50% in 75 mM HEPES, pH 8.0) (Boehringer-Mannheim) is stirred into the pellet, followed by the addition of serum-free RPMI. Thereafter, the cells are centrifuged and resuspended in RPMI containing 15% FBS, 100 μM sodium hypoxanthine, 0.4 μM aminopterin, 16 μM thymidine (HAT) (Gibco), 25 Units/ml IL-6 (Boehringer-Mannheim) and 1.5 x 10^6 thymocytes/ml and plated into 10 Corning flat-bottom 96-well tissue culture plates (Corning, Corning New York).

On days 2, 4, and 6 after the fusion, 100 μl of medium is removed from the wells of the fusion plates and replaced with fresh medium. On day 8, the fusions are screened by ELISA, testing for the presence of mouse IgG that binds to a GPCR-like receptor polypeptide. Selected fusions are further cloned by dilution until monoclonal cultures producing anti-GPCR-like receptor antibodies are obtained.

B. GPCR-like Receptor-Neutralizing Antibodies from Phage Display

GPCR-like receptor-neutralizing antibodies are generated by phage display techniques such as those described in Aujame *et al., Human Antibodies, 8(4)*:155-168 (1997); Hoogenboom, *TIBTECH, 15*:62-70 (1997); and Rader *et al., Curr. Opin. Biotechnol., 8*:503-508 (1997), all of which are incorporated by reference. For example, antibody variable regions in the form of Fab fragments or linked single chain Fv fragments are fused to the amino terminus of filamentous phage minor coat protein pIII. Expression of the fusion protein and incorporation thereof into the mature phage coat results in phage particles that present an antibody on their surface and contain the genetic material encoding the antibody. A phage library comprising such constructs is expressed in bacteria, and the library is screened for GPCR-like receptor-specific phage-antibodies using labeled or immobilized GPCR-like receptor as antigen-probe.

C. GPCR-like Receptor-Neutralizing Antibodies from Transgenic Animals

GPCR-like receptor-neutralizing antibodies are generated in transgenic animals, such as mice, essentially as described in Bruggemann *et al., Immunol. Today*

17(8):391-97 (1996) and Bruggemann *et al.*, *Curr. Opin. Biotechnol. 8*:455-58 (1997). Transgenic mice carrying V-gene segments in germline configuration, and expressing the transgenes in their lymphoid tissue, are immunized with a GPCR-like polypeptide composition using conventional immunization protocols. Hybridomas are generated from B cells of the immunized mice using conventional protocols and screened to identify hybridomas secreting anti-GPCR-like receptor antibodies (*e.g.*, as described above).

Example 5

Assays to Identify Modulators of GPCR-Like Receptor Activity

Two general approaches exist for the discovery of novel compounds that are binding partners for receptors, especially GPCRs. Each of the general approaches is compatible with high throughput screening (*i.e.*, HTS) formats, which are preferred formats for identifying ligands and other binding partners of GPCR-like receptor polypeptides, such as modulators of receptor activity. The first approach involves measuring the binding of a known ligand, preferably labeled with a radiolabel, to a preparation that contains the receptor, either found in native tissue or based on expression of the gene encoding the receptor, typically in a heterologous system. The recombinant system involves the expression of a recombinant GPCR, and is the presently preferred system for binding assays. The second approach to the identification of GPCR-like receptor binding partners involves the measurement of the activity of the receptor, which may be influenced by either the binding of ligands that are agonists, and elevate receptor activity, or by the binding of antagonists, which interfere with agonist binding, thereby reducing the level of receptor activity. As for binding assays, recombinant systems are the presently preferred forms for function-based assays.

Receptor binding assays (RBA) as HTS systems for drug discovery

The literature is replete with examples of the use of radiolabeled ligands in HTS binding assays for drug discovery (see Williams, Med. Res. Rev. 11:147-184 (1991); Sweetnam *et al.*, J. Nat. Prod. 56:441-455 (1993) for review). It is also possible to screen for novel anthelmintic compounds with radiolabeled ligands in HTS binding screens (Geary et al., 1999). However, native nematode tissue, particularly tissue from parasitic nematodes, is difficult to acquire and native nematode tissue binding assays are not well-suited for HTS because the density of receptor sites in these preparations is typically much lower than found in mammalian tissues (Thompson, *et al.*, Parasitology 113:S217-S238 (1996)). Other reasons that recombinant receptors are preferred for HTS binding assays include better specificity (higher relative purity) and ability to generate large amounts of receptor material (see Hodgson, Bio/Technology 10:973-980 (1992)).

A variety of heterologous systems are available for expression of recombinant receptors and are well known to those skilled in the art. Such systems include bacteria (Strosberg *et al.*, Trends in Pharm. Sci. 13:95-98 (1992)), yeast (Pausch, Trends in Biotech. 15:487-494 (1997)), several kinds of insect cells (Vanden Broeck, Intl. Rev. Cytol. 164:189-268 (1996)), amphibian cells (Jayawickreme *et al.*, Curr. Opin. Biotechnol. 8:629-634 (1997)) and several mammalian cell lines (CHO, HEK293, COS, etc.; see Gerhardt *et al.*, Eur. J. Pharmacol. 334:1-23 (1997); Wilson *et al.*, Brit. J. Pharmacol. 125:1387-1392 (1998)). These examples do not preclude the use of other possible cell expression systems, including cell lines obtained from nematodes (WO 98/37177).

A *C. elegans* orphan GPCR expressed in one of the described recombinant systems can be used for HTS binding assays in conjunction with its defined ligand, in this case the corresponding neuropeptide (*e.g.*, FaRP) that activates it. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, ^{125}I (preferred; see Geary et al., 1999), ^{3}H, ^{35}S or ^{32}P, by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur *et al.*, Drug Dev. Res. 33:373-398 (1994); Rogers, Drug Disc. Today 2:156-160 (1997)). Radioactive ligand specifically bound to the receptor in membrane preparations made from the cell line expressing the

recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate bound ligand from unbound ligand (Williams, 1991; Sweetnam *et al.*, 1993). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama *et al.*, Drug Disc. & Dev. 1:85-91 (1998); Boss *et al.*, J. Biomol. Screening 3:285-292 (1998)). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (see Rogers, 1997; Hill, Curr. Opin. in Drug Disc. & Dev. 1:92-97 (1998)).

A binding assay that has proven useful is a Flp18 binding assay or screen.

A peptide ligand screening assay that measures ligand binding to a suitable GPCR-like receptor according to the invention is exemplified using a modified *C. elegans* Flp18 peptide ligand. The modified *C. elegans* Flp18 peptide, YDVPGVLRFamide (SEQ ID NO:147), is extended by an N-terminal tyrosine residue. Iodination of the peptide is accomplished using a standard chloramine T procedure. Added to a 2 ml glass vial are 10 µl of a 1 mM water solution of the Flp18 peptide, 10 µl of 0.1M (pH 7.99) sodium phosphate buffer, 1.0 mCi [^{125}I] sodium iodide and 5 µl of a 2 mg/ml chloramine T solution (in the phosphate buffer). The mixture is vortexed for 60 seconds and the reaction is stopped by the addition of 25 µl of a 5 mg/ml solution of sodium metabisulfite in phosphate buffer. The mixture is then subjected to HPLC gradient fractionation by injecting it onto a Vydac C18 (0.45 x 15 cm) column. The gradient used is 70% A and 30% B at time zero to 20% A and 80% B at time 25 minutes (A = 0.1M NH_4 acetate in water; B = 0.1M NH_4 acetate in water 40%: CH_3CN 60%, v/v). The flow rate is 1.0 ml/minute. Samples are collected into 0.25 ml capture buffer (O.1M sodium phosphate buffer with 0.5% bovine serum albumin, 0.1% Triton X100 and 0.05% Tween 20) at 30 second intervals from t = 8 to t = 20 minutes. Monoiodo YDVPGVLRFamide (SEQ ID NO:119) typically elutes at t = 11 minutes and the yield is approximately 100 µCi in 0.75 ml.

Membranes for use in the assay are derived from cells stably transfected with a GPCR-like receptor expression construct. The transfected cells are grown in 10 cm dishes to subconfluence, rinsed once with 5 ml of ice-cold Ca^{2+}/Mg^{2+}- free PBS, and

scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x *g*, 5 minutes), resuspended in 25 mM Tris, pH 7.5, 5 mM EDTA, 5 mM EDTA, pH 7.5 (TEE), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (one ml TEE per plate of cells), and centrifuged at 1,000 x *g* for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x *g* for 20 minutes to isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA). The resuspended membranes are optionally frozen in liquid nitrogen and stored at -70°C until use.

Aliquots of cell membranes prepared as described above and stored at -70°C are thawed, homogenized, and diluted into assay buffer as described below.

The peptide ligand screening assay is facilitated by use of 96-well plates Millipore Multiscreen® filtration plates (FB opaque 1.0 μM glass fiber type B, catalog number MAFBNOB50). Protein reaction products are collected using a Millipore Multiscreen® solvent resistant manifold (catalog number MAVMO960R). Each replicate of the assay is contained in a single well and has a volume of 100 ul containing 5 ug protein (preparation described above). Each test group contains two replicates. For each test compound, one group is run with [^{125}I]YDVPGVLRFamide (SEQ ID NO:119) only to determine total binding, and one group is run with 1 μM (or as designated) concentration of the test compound and [^{125}I]YDVPGVLRFamide (SEQ ID NO:119; for non-specific binding). The order of adding reagents for each replicate is as follows: assay buffer (20 mM HEPES, 10 mM $MgCl_2$, 1% bovine serum albumin, pH 7.4) test compound (made up in assay buffer), [^{125}I]YDVPGVLRFamide (SEQ ID NO:119; in assay buffer) and membrane suspension (in assay buffer). The addition of the membrane suspension initiates the binding reaction which is run for 30 minutes at room temperature (22°C). Following the 30 minute incubation each plate is placed on the filtration manifold and vacuum is applied, pulling the liquid through the filter (the liquid is discarded) and catching the protein on the filters in each well. For washing, the vacuum is released and 200 μl assay buffer is added to each well followed by reapplication of the vacuum. This washing is repeated twice more (total of 3 washes for each replicate).

Following washing, the plastic covering on the underside of each plate is removed and the plate is placed in a bottom-sealed Microbeta® scintillation counting cassette (catalog number 1450-105). Twenty five µl of scintillant is added to each well and the plate is placed on a rotary shaker at 80 rpm for one hour and then allowed to sit overnight. The following day, the plate is counted in a Microbeta® scintillation counter. The mean non-specific binding is subtracted from the mean total binding to yield specific binding for both the standard (YDVPGVLRFamide, SEQ ID NO:147) and the unknowns.

Response-based GPCR HTS systems

It is well known that activation of heterologous receptors expressed in recombinant systems results in a variety of biological responses, which are typically mediated by G proteins expressed in the host cells. Agonist binding to a GPCR results in exchange of bound GDP for GTP at a binding site on the G subunit; one can use a radioactive, non-hydrolyzable derivative of GTP, such as [^{35}S]GTPγS, to measure binding of an agonist to the receptor. (Seifert *et al.*, Eur. J. Biochem. 255:369-382 (1998).) One can also use this binding to measure the ability of antagonists to bind to the receptor by decreasing binding of GTP [^{35}S] in the presence of a known agonist.

The G proteins required for functional expression of heterologous GPCRs can be native constituents of the host cell or can be introduced through well-known recombinant technology. The G proteins can be intact or chimeric. Often, a nearly universally competent G protein (*e.g.*, $G_{\alpha 16}$) is used to couple any given receptor to a detectable response pathway. G protein activation results in the stimulation or inhibition of other native proteins, events that can be linked to a measurable response.

Examples of such biological responses include, but are not limited to, the following responses: the ability to survive in the absence of a limiting nutrient in specifically engineered yeast cells (Pausch, 1997); changes in intracellular Ca^{2+} concentration as measured by fluorescent dyes (Murphy *et al.*, Curr. Opin. in Drug Disc. & Dev. 1:192-199 (1998)). Fluorescence changes can also be used to monitor ligand-induced changes in membrane potential or intracellular pH; an automated system suitable for HTS has been described for these purposes (Schroeder *et al.*, J. Biomol. Screening 1:75-80 (1996)). Melanophores prepared from *Xenopus laevis* show a ligand-dependent

change in pigment organization in response to heterologous GPCR activation; this response is adaptable to HTS formats (Jayawickreme *et al.*, 1997). Assays are also available for the measurement of common second messengers, including cAMP, phosphoinositides and arachidonic acid. Set forth in the following subsections are exemplary function-based assays for identifying modulators (agonists and antagonists) of GPCR-like receptor activity. Among the modulators that can be identified by these assays are natural ligand compounds of the receptor; synthetic analogs and derivatives of natural ligands; antibodies, antibody fragments, and/or antibody-like compounds derived from natural antibodies or from antibody-like combinatorial libraries; and/or synthetic compounds identified by high throughput screening of libraries; and other libraries known in the art. All modulators that bind GPCR-like receptors are useful for identifying GPCR-like polypeptides in tissue samples (*e.g.*, for diagnostic purposes, pathological purposes, and other purposes known in the art). Agonist and antagonist modulators are useful for up-regulating and down-regulating GPCR-like receptor activity, respectively, to treat invertebrate infestations and/or infections of man, other animals, plants or the environment. GPCR-like receptor binding partners also may be used to deliver a therapeutic compound or a label to cells that express a GPCR-like receptor (*e.g.*, by attaching the compound or label to the binding partner). The assays may be performed using single putative modulators; they may also be performed using a known agonist in combination with candidate antagonists (or *visa versa*).

A. cAMP Assays

In one type of assay, levels of cyclic adenosine monophosphate (cAMP) are measured in GPCR-like receptor-transfected cells that have been exposed to candidate modulator compounds. Protocols for cAMP assays have been described in the literature. [See, *e.g.*, Sutherland *et al.*, Circulation, 37:279 (1968); Frandsen, *et al.*, Life Sciences, 18:529-541 (1976); Dooley *et al.*, J. Pharm. & Exp. Therap., 283(2):735-41 (1997); and George *et al.*, J. Biomol. Screening, 2(4):235-40 (1997).] An exemplary protocol for such an assay, using an Adenylyl Cyclase Activation FlashPlate® Assay from NEN™ Life Science Products, is set forth below.

Briefly, the GPCR-like receptor coding sequence (*e.g.*, a cDNA or intronless genomic DNA) is subcloned into a commercial expression vector, such as pzeoSV2 (Invitrogen, San Diego, CA), and transiently transfected into Chinese Hamster Ovary (CHO) cells using known methods, such as the transfection reagent FuGENE 6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert. The transfected CHO cells are seeded into the 96-well microplates of the FlashPlate® assay kit, which are coated with solid scintillant to which antisera to cAMP has been bound. For a control, some wells are seeded with wild type (untransfected) CHO cells. Other wells on the plate receive various amounts of cAMP standard solution for use in creating a standard curve.

One or more test compounds are added to the cells in each well, with water and/or compound-free medium/diluent serving as a control. After treatment, cAMP is allowed to accumulate in the cells for exactly 15 minutes at room temperature. The assay is terminated by the addition of lysis buffer containing [^{125}I]-labeled cAMP, and the plate is counted using a Packard Topcount™ 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells (or from standards) competes with the fixed amounts of [^{125}I]-cAMP for antibody bound to the plate. A standard curve is constructed, and cAMP values for the unknowns are obtained by interpolation. Changes in intracellular cAMP level of the cells in response to exposure to a test compound are indicative of GPCR-like receptor modulating activity. Modulators that act as agonists at receptors which couple to the G_s subtype of G proteins will stimulate production of cAMP, leading to a measurable 3-10 fold increase. Agonists OF receptors which couple to the $G_{i/o}$ subtype of G proteins will inhibit forskolin-stimulated cAMP production, leading to a measurable decrease of 50-100%. Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

B. Aequorin Assays

In another function based assay, cells (*e.g.*, CHO cells) are transiently co-transfected with both a GPCR-like receptor expression construct and a construct that encodes the photoprotein apoaquorin. In the presence of the cofactor coelenterazine,

apoaequorin will emit a measurable luminescence that is proportional to the amount of intracellular (cytoplasmic) free calcium. [See generally Cobbold P.H. and Lee, J.A.C. "Aequorin measurements of cytoplasmic free calcium. In: McCormack J.G. and Cobbold P.H., eds., *Cellular Calcium: A Practical Approach*. Oxford:IRL Press (1991); Stables *et al., Analytical Biochemistry, 252*: 115-26 (1997); and Haugland, R.P. *Handbook of Fluorescent Probes and Research Chemicals*. Sixth edition. Eugene OR: Molecular Probes (1996).]

In one exemplary assay, a GPCR-like receptor is subcloned into the commercial expression vector pzeoSV2 (Invitrogen, San Diego, CA) and transiently co-transfected along with a construct that encodes the photoprotein apoaequorin (Molecular Probes, Eugene, OR) into CHO cells using the transfection reagent FuGENE 6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert. Alternatively, the GPCR-like receptor may be subcloned into pCR3.1 and transfected into CHO cells using LipofectAMINE PLUS™, as described above.

The cells are cultured for 24 hours at 37°C in DMEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 Units/ml penicillin and 10 µg/ml streptomycin, at which time the medium is changed to serum-free DMEM containing 5 µM coelenterazine (Molecular Probes, Eugene, OR), and the cells are cultured for two additional hours at 37°C. Cells are then detached from the plate using VERSEN (Gibco/BRL), washed, and resuspended at 200,000 cells/ml in serum-free DMEM.

Dilutions of candidate GPCR-like receptor modulator drugs are prepared in serum-free DMEM and dispensed into wells of an opaque 96-well assay plate, 50 µl/well. Plates are loaded onto an MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantilly, VA). The instrument is programmed to dispense a 50 µl cell suspension into each well, one well at a time, and immediately read luminescence for 15 seconds. Dose-response curves for the modulator candidates are constructed using the area under the curve for each light signal peak. Data are analyzed with SlideWrite, using the equation for 1-site ligand, and EC_{50} values are obtained. Changes in luminescence caused by the drugs are considered indicative of modulatory activity. Modulators that act as agonists at receptors which couple to the G_q subtype of G proteins give an increase in

luminescence of up to 100 fold. Modulators that act as inverse agonists will reverse this effect at receptors that are either constitutively active or activated by known agonists.

C. Luciferase Reporter Gene Assay

The photoprotein luciferase provides another useful tool for assaying for modulators of GPCR-like receptor activity. Cells (*e.g.*, CHO cells or COS 7 cells) are transiently co-transfected with both a GPCR-like receptor expression construct (*e.g.*, a GPCR-like receptor in pzeoSV2 (Invitrogen, San Diego, CA)) and a reporter construct which includes a luciferase coding region downstream from a transcription factor, either the cAMP-response element (CRE), AP-1, or NF-κB. Agonist binding to receptors coupled to the G_s subtype of G proteins leads to increases in cAMP, activating the CRE transcription factor and resulting in expression of the luciferase gene. Agonist binding to receptors coupled to the G_q subtype of G proteins leads to production of diacylglycerol that activates protein kinase C, which activates the AP-1 or NF-κB transcription factors resulting in expression of the luciferase gene. Expression levels of luciferase reflect the activation status of the signaling events. [See generally George *et al., Journal of Biomolecular Screening, 2(4)*: 235-40 (1997); and Stratowa *et al., Current Opinion in Biotechnology, 6*: 574-81 (1995).] Luciferase activity may be quantitatively measured using, *e.g.*, luciferase assay reagents that are commercially available from Promega (Madison, WI).

In one exemplary assay, CHO cells are plated in 24-well culture dishes at a density of 100,000 cells/well one day prior to transfection and cultured at 37°C in DMEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 Units/ml penicillin and 10 μg/ml streptomycin. Cells are transiently co-transfected with both a GPCR-like receptor expression construct and a reporter construct containing the luciferase gene. The reporter plasmids CRE-luciferase, AP-1-luciferase and NF-κB-luciferase may be purchased from Stratagene (LaJolla, CA). Transfections are performed using FuGENE 6 transfection reagent (Boehringer-Mannheim), following the protocol provided in the product insert. Cells transfected with the reporter construct alone are used as a control. Twenty-four hours after transfection, cells are washed once with phosphate buffered saline (PBS)

pre-warmed to 37°C. Serum-free DMEM is then added to the cells either alone (control) or with one or more candidate modulators and the cells are incubated at 37°C for five hours. Thereafter, cells are washed once with ice-cold PBS and lysed by the addition of 100 µl of lysis buffer/well (from the luciferase assay kit, Promega, Madison, WI). After incubation for 15 minutes at room temperature, 15 µl of the lysate is mixed with 50 µl substrate solution (Promega) in an opaque white 96-well plate, and the luminescence is read immediately on a Wallace model 1450 MicroBeta scintillation and luminescence counter (Wallace Instruments, Gaithersburg, MD).

Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulating activity. Receptors that are either constitutively active or activated by agonists typically give a 3-20 fold stimulation of luminescence compared to cells transfected with the reporter gene alone. Modulators that act as inverse agonists will reverse this effect.

D. Intracellular calcium measurement using FLIPR

Changes in intracellular calcium levels are another recognized indicator of G protein-coupled receptor activity, and such assays can be employed to assay for modulators of GPCR-like receptor activity. For example, CHO cells stably transfected with a GPCR-like receptor expression construct are plated at a density of 4×10^4 cells/well in Packard black-walled 96-well plates specially designed to isolate fluorescent signals to individual wells. The cells are incubated for 60 minutes at 37°C in modified Dulbecco's PBS (D-PBS), containing 36 mg/L pyruvate and 1 g/L glucose with the addition of 2.5 mM probenecid (Sigma Chemical Co.) and one of four calcium indicator dyes (Fluo-3™ AM, Fluo-4™ AM, Calcium Green™-1 AM, or Oregon Green™ 488 BAPTA-1 AM) at a concentration of 4 µM. Plates are washed once with modified D-PBS without 1% fetal bovine serum and incubated for 10 minutes at 37°C to remove residual dye from the cellular membrane. In addition, a series of washes with modified D-PBS is performed immediately prior to activation of the calcium response.

Calcium response is initiated by the addition of one or more candidate receptor agonist compounds, or a positive control such as a calcium ionophore A23187 (10 µM), or ATP (4 µM). Fluorescence is measured by Molecular Device's FLIPR with

an argon laser, excitation at 488 nm. [See, *e.g.*, Kuntzweiler *et al., Drug Development Research, 44(1)*: 14-20 (1998).] The F-stop for the detector camera was set at 2.5 and the length of exposure was 0.4 msec. Basal fluorescence of cells was measured for 20 seconds prior to addition of agonist, ATP, or A23187, and was subtracted from the response signal. The calcium signal is measured for approximately 200 seconds, taking readings every two seconds. The calcium ionophore and ATP increase the calcium signal 200% above baseline levels. In general, activated GPCRs increase the calcium signal approximately 10-15% above baseline signal.

E. Mitogenesis Assay

In mitogenesis assays, the ability of candidate modulators to induce or inhibit GPCR-like receptor-mediated cell growth is determined. [See, e.g., Lajiness *et al., Journal of Pharmacology and Experimental Therapeutics, 267(3)*: 1573-81 (1993).] For example, CHO cells stably expressing a GPCR-like receptor are seeded into 96-well plates at a density of 5000 cells/well and grown at 37°C in DMEM with 10% fetal calf serum for 48 hours, at which time the cells are rinsed twice with serum-free DMEM. After rinsing, 80 µl of fresh DMEM, or DMEM containing a known mitogen, is added along with 20 µl DMEM containing varying concentrations of one or more test compounds diluted in serum-free medium. As controls, some wells on each plate receive serum-free medium alone, and some receive medium containing 10% fetal bovine serum. Untransfected cells or cells transfected with the vector alone also may serve as controls.

After culture for 16-18 hours, 1 µCi/well of [^{3}H]-thymidine is added to the wells and cells are incubated for an additional 2 hours at 37°C. The cells are trypsinized and harvested onto filter mats with a cell harvester (Tomtec) and the filters are counted in a Betaplate counter. The incorporation of [^{3}H]-thymidine in serum-free test wells is compared to the results achieved in cells stimulated with serum. Use of multiple concentrations of test compounds permits creation and analysis of dose-response curves using the non-linear, least squares fit equation: $A = B \times [C/ (D + C)] + G$ where A is the percent of serum stimulation; B is the maximal effect minus baseline; C is the EC_{50}; D is the concentration of the compound; and G is the maximal effect. Parameters B, C and G are determined by Simplex optimization.

Agonists that bind to the receptor are expected to increase [^{3}H]-thymidine incorporation into cells, showing up to 80% of the response to serum. Antagonists that bind to the receptor will inhibit the stimulation seen with a known agonist by up to 100%.

F. [^{35}S]GTPγS Binding Assay

Because G protein-coupled receptors signal through intracellular G proteins whose activity involves GTP/GDP binding and hydrolysis, measurement of binding of the non-hydrolyzable GTP analog [^{35}S]GTPγS in the presence and absence of putative modulators provides another indicator of modulator activity. [See, *e.g.*, Kowal, *et al., Neuropharmacology, 37*: 179-87 (1998).]

In one exemplary assay, cells stably transfected with a GPCR-like receptor expression construct are grown in 10 cm dishes to subconfluence, rinsed once with 5 ml of ice-cold Ca^{2+}/Mg^{2+}- free PBS, and scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x *g*, 5 minutes), resuspended in 25 mM Tris, pH 7.5, 5 mM EDTA, 5 mM EDTA, pH 7.5 (TEE), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (one ml TEE per plate of cells), and centrifuged at 1,000 x *g* for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x *g* for 20 minutes to isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at -70°C until use.

Aliquots of cell membranes prepared as described above and stored at -70°C were thawed, homogenized, and diluted into buffer containing 20 mM HEPES, pH 7.4, 10 mM $MgCl_2$, 1 mM EDTA, and 100 mM NaCl at a concentration of 10-50 μg/ml. Reaction mixtures were prepared in 96-well polypropylene plate, Nunc. Twenty microliters of a 10X putative modulator compound (including but not limited to potential peptide ligands) solution or a water control (20 μl), 18.2 M GDP (0.11 ml, 10 μM final concentration), and a membrane preparation were mixed and placed on ice. The modulator/ligand-GDP-membrane mixtures were incubated for 20 minutes at room

temperature on a shaking platform and then placed on ice. To each sample, 20 µl guanosine 5'-O-(3[^{35}S]thio) triphosphate (NEN, 600-1200 Ci/mmol; [^{35}S]-GTPγS), was added to approximately 40,000 cpm/0.2 ml, or a final concentration of 100 pM. Plates with the incubation mixtures (0.2 ml/well total) were incubated at room temperature for 45 minutes. Reaction mixture aliquots, 0.175 ml each, were then transferred into pretreated (100 µl/well wash buffer) 96-well FB MultiScreen filter plates (Millipore). Membranes were subsequently washed three times with 0.25 ml ice-cold wash buffer (10 mM HEPES, pH 7.4, 10 mM $MgCl_2$) per well each time and vacuum-filtered. After the last wash, Supermix Optiphase scintillation fluid (25 µl/well, Wallac) was added and the plates were sealed and counted in a Trilux 1450 Microbeta counter (Wallac) for one minute per well. As positive controls, membranes from CHO cells stably expressing a rat dopamine type 2 (rD_2) receptor, were treated with 1 mM dopamine in 0.025% ascorbic acid (100 µM dopamine final concentration). Non-specific binding was measured in the presence of 100 µM cold GTPγS and was subtracted from the total. Each treatment was typically carried out in triplicate.

Ligand-induced stimulation of [^{35}S]GTPγS binding was expressed as a multiple of the basal activity with no ligand added. Each treatment was run either in triplicate, or, on occasion in duplicate and bindings (cpm) were calculated as means +/- standard deviations. Dose-response curves for the receptor/ligand systems were analyzed using a non-linear least squares SAS model, with $y= B_{max} X/(K_d + X)$. Other dose-response curves were analyzed using Prism (GraphPad Software, Inc. San Diego, CA) and the following equation $y = \text{Bottom} + (\text{Top} - \text{Bottom})/ (1 + 10^{\text{LogEC}_{50}-X})$.

Originally, the GTPγS assay was chosen as a functional assay because agonist-driven stimulation revealed by the GTPγS assay reflects early events in the GPCR activation cascade, regardless of further activation pathways of various down-stream signaling events. This appears especially useful for the assessment of possible activation of orphan GPCRs (*i.e.*, GPCRs having unknown functions and unknown signaling pathways). The GTPγS assay was carried out with membranes prepared from CHO cells transiently transfected with DNA encoding *C. elegans* GPCRs using a 96-well MultiScreen G/FB filter plates and a MultiScreen vacuum manifold (Millipore) for filtration. Because the GTPγS assay is known to poorly recognize GPCRs coupled to the

Gq class of G-proteins, a Ca^{+2} mobilization assay, based on a FLIPR readout, was used as well to evaluate putative Gq-coupled orphan GPCRs transiently transfected into CHO cells. It was essential to use a 37°C to 28°C temperature downshift and to incubate the cells for an additional 24-48 hours for functional receptor expression. The temperature shift was implemented 24 hours after transfection. Without the temperature downshift, expression of functional receptor could not be detected.

The CEGPCR3 receptor was found to be activated by several peptide ligands (Table 7), as determined in the GTPγS assay. Chinese hamster ovary cells were incubated for 24 hours at 37°C after transfection, followed by an additional 24 hours incubation at 28°C before cell harvesting for membrane preparation. CEGPCR3 was matched with two *C. elegans* peptides encoded by *flp15*, GGPQGPLRF-NH_2 (EC_{50} 152 nM) and GPSGPLRF-NH_2 (EC_{50} 422 nM). A *Manduca* peptide, GNSFLRF-NH_2 (EC_{50} 7900 nM), also activated the CEGPCR3 receptor, albeit with a potency about 19-52-fold lower than that determined for the two *C. elegans flp15* peptides. Based on these data, we identified CEGPCR3 as the receptor for *flp15* peptides.

Table 7

Receptor	Peptide ligand	SEQ ID NO	Species of origin	EC_{50} range
CEGPCR3 (SEQ ID NO:44)	GGPQGPLRF-NH_2	85	*C. elegans/flp15*	
	GPSGPLRF-NH_2	89	*C. elegans/flp15*	150 –400 nM
	GNSFLRF-NH_2	88	*Manduca*	~ 8 μM
CEGPCR7 (SEQ ID NO:26)	GLGPRPLRF-NH_2	86	*A. suum (AF9)/ C.elegans (flp21)*	
	[I]Y^0-GLGPRPLRF-NH_2	118	Synthetic AF9 analog	~200-250 nM

In addition, peptide ligands were identified for the CEGPCR7 receptor, as revealed in Table 7. One peptide bears the sequence GLGPRPLRF-NH_2 (AF9) (EC_{50} 207 nM). It is worth noting that [I]Y^0GLGPRPLRF-NH_2, representing an AF9 analog N-terminally extended with a 3-iodo-Tyr residue, was also active (EC_{50} 237 nM). The

functional activity of [I]Y^0AF9 indicates that this analog is an agonist and, in labeled (*e.g.*, radioiodinated) form is useful as a probe for binding assays, including high-throughput screening (HTS) assays.

The CEGPCR4 receptor was activated by several peptide ligands, as determined using the GTPγS assay (Table 8). To perform the assays, CHO cells were incubated as described above in the context of assaying CEGPCR3. The *C. elegans* peptides encoded on *flp18* were the most potent (EC_{50}'s in the low nM range). Also very potent were several *A. suum* peptides, with high sequence homology to the *C. elegans* FLP18 peptides (all having the same sequence, PGVLRF-NH_2, at their C-termini). In addition to the FLP18 FaRPs, a battery of peptides from different invertebrate species *(Manduca, lobster, locust, Drosophila*), also activated this receptor. The latter were, however, significantly less potent (EC_{50}s in the high nanomolar to micromolar range). Based on the EC_{50} values, we identified CEGPCR4 as the receptor for *C. elegans flp18* peptides. As shown in Table 8, [IodoY]DVPGVLRF-NH_2, which is a N-terminally iodo-Tyr-extended *flp18* analog (SEQ ID NO:119, EC_{50} = 8), was essentially as potent as its parent peptide, DVPGVLRF-NH_2. This finding, as well as high affinity binding of [^{125}I-Y]DVPGVLRF-NH_2 (SEQ ID NO:80) to the CEGPCR4/CHO membranes (Kd in the subnanomolar range), showed that [^{125}I-Y]DVPGVLRF-NH_2 (SEQ ID NO:119) is an excellent probe, preferably radioiodinated, for a high throughput binding assay as a screen for modulators of the ligand-receptor interaction.

Shown in Table 8 are the synthetic peptides, various analogs of DVPGVLRF-NH_2 that acted as agonists of the CEGPCR4 receptor in CHO cells. Most of the synthetic peptides were very active (EC_{50}s in the nanomolar range in the GTPγS assay, and IC_{50}s in the low nanomolar range in the competition binding assay). The C-terminal tetrapeptide VLRF-NH2 was considerably less potent with an EC_{50} of about 4 micromolar in the GTPγS assay, and an IC50 of about 400 nM in the competition binding assay. In general, iodo-Tyr-modified peptide analogs, particularly those disclosed herein (*e.g.*, Table 8), represent good potential probes for binding assays and may be conveniently labeled with ^{125}I..

Table 8

Receptor	Peptide ligand	SEQ ID NO	Species of origin	EC_{50} range-GTPγS assay	IC_{50} range-binding assay
CEGPCR4 (SEQ ID NO:22)	DVPGVLRF-NH_2	80	*C. elegans/flp18*		
	KSVPGVLRF-NH_2	92	*C. elegans/flp18*		
	SEVPGVLRF-NH_2	98	*C. elegans/flp18*		
	SVPGVLRF-NH_2	100	*C. elegans/flp18*		
	DFDGAMPGVLRF-NH_2	120	*C. elegans/flp18*		
	EIPGVLRF-NH_2	171	*C. elegans/flp18*	~5-80 nM	0.5-10 nM
	AVPGVLRF-NH_2 (AF3)	79	*A. suum*		
	GDVPGVLRF-NH_2 (AF4)	84	*A. suum*		
	GMPGVLRF-NH_2 (AF20)	87	*A. suum*		
	ASPSFIRF-NH_2	78	*C. elegans/flp4*		
CEGPCR4 (SEQ ID NO:22)	GNSFLRF-NH_2	88	*Manduca*		
	KPNFLRF-NH_2	91	*C. elegans/flp1*		
	PDVDHVFLRF-NH_2 (SchistoFLRFa)	94	*Locust*	~0.4 - 9 μM	60-900 μM
	pQDVDHVFLRF-NH_2 (leucomyosuppressin)#	95	*Locust*		
	ILNleRF-NH_2	90	synthetic		
CEGPCR4 (SEQ ID NO:22)	SPLGTMRF-NH_2	143	*C. elegans/flp3*		
	SDNFMRF-NH_2	122	Drosophila		
	PDNFMRF-NH_2	123	Drosophila		
	SAEPFGTMRF-NH_2	97	*C. elegans/flp3*		
	GGPQGPLRF-NH_2	85	*C. elegans/flp15*	~10 μM or higher	50-500 μM
	EIVFHQISPIFFRF-NH_2	83	*C. elegans/flp14*		
	TDVDHVFLRF-NH_2	101	*Drosophila*		
	TNRNFLRF-NH_2 (Lobster peptide II)	102	*Lobster*		
	NGAPQPFVRF-NH_2	93	*C. elegans/flp11*		
CEGPCR4 (SEQ ID NO:22)	VLRF-NH_2	152	synthetic	~4 μM	~0.4 μM

Receptor	Peptide ligand	SEQ ID NO	Species of origin	EC_{50} range-GTPγS assay	IC_{50} range-binding assay
CEGPCR4 (SEQ ID NO:22)	GVLRF-NH_2	153	Synthetic peptides	~5 - 400 nM	~1-60 nM
	Ac-GVLRF-NH_2	154			
	PGVLRF-NH_2	155			
	VPGVLRF-NH_2	156			
	Ac-VPGVLRF-NH_2	157			
	DVPGVLRF-NH_2	80			
	Y-DVPGVLRF-NH_2	158			
	3-Iodo-Y-DVPGVLRF-NH_2	119			
	Y-VPGVLRF-NH_2	159			
	3-Iodo-Y-VPGVLRF-NH_2	160			
	Y-PGVLRF-NH_2	161			
	3-Iodo-Y-PGVLRF-NH_2	162			
	Y-GVLRF-NH_2	163			
	3-Iodo-Y-GVLRF-NH_2	164			

Evaluation of CEGPCR5 expressed in CHO cells in the [^{35}S]GTPγS assay, as described above, yielded several active peptide ligands, as shown in Table 9. A peptide ligand designated AF9 (which also binds CEGPCR7, see above) bearing the sequence GLGPRPLRF-NH_2 (SEQ ID NO:86) was the most potent activator (EC_{50} in the low nanomolar range) of CEGPCR5. An analog extending AF9 by an iodo-Tyr residue at its N-terminus, [iodoY]-GLGPRPLRF-NH_2 (SEQ ID NO:118), was about 200-fold less potent (EC_{50} range of 176-273 nM), but was active enough to serve as a ligand in binding assays designed to identify modulators. Based on the GTPγS assay results, the *C. elegans* CEGPCR5 was identified as an AF9 or *flp21* peptide receptor.

Table 9

Receptor	Peptide	SEQ ID NO	Species of origin	EC_{50} range
CEGPCR5 (SEQ ID NO:46)	GLGPRPLRF-NH_2 (AF9)	86	*C. elegans(flp21), A. suum*	~1-10 nM
CEGPCR5 (SEQ ID NO:46)	YGLGPRPLRF-NH_2	125	*[Y^0]AF9 analog*	
	[iodoY]GLGPRPLRF-NH_2	118	*[iodoY^0]AF9 analog*	
	GNSFLRF-NH_2	88	*Manducca*	
	YLRF-NH_2	126	Leech	
	SDRNFLRF-NH_2	127	Lobster	
	FLRF-NH_2	128	*Trematoda*	~80- 800 nM
	FMRF-NH_2	129	*Mollusca*	
	TNRNFLRF-NH_2	102	Lobster	
	DPSFLRF-NH_2	131	*Manducca*	
	SVPGVLRF-NH_2	100	*C. elegans (flp18)*	
CEGPCR5 (SEQ ID NO:46)	AGPRFIRF-NH_2	133	*A suum*	
	KSVPGVLRF-NH_2	92	*C. elegans (flp18)*	
	KPNFLRY-NH_2	135	*C. elegans (flp1)*	
	ARGPQLRLRF-NH_2	136	*L. decemlineata*	
	YIRF-NH_2	137	*Trematoda*	
	GMPGVLRF-NH_2	87	*A. suum*	
	AGAKFIRF-NH_2	139	*C. elegans (flp5)*	~1 - 8 μM
	KPNFLRF-NH_2	91	*C. elegans (flp1)*	
	KHEYLRF-NH_2	141	*C. elegans (flp14)*	
	EIVFHQISPIFFRF-NH_2	83	*C. elegans (flp14)*	

The CEGPCR16 receptor, expressed in CHO cells as described for CEGPRC3, was also used in a [^{35}S]GTPγS assay to screen for peptide ligands, which resulted in the identification of peptides encoded by the *C. elegans flp3* gene. Of the peptides investigated, the Flp3 peptides were the most potent activators (Table 10), with EC_{50} values of 100-330 nM. One noteworthy ligand is the Flp18 peptide, SVPGVLRF-NH_2 (SEQ ID NO:100, EC_{50} = 64 nM), which is comparable to, or more potent than, the

Flp3 peptides in activating CEGPCR16; most of the other Flp18 peptides were either inactive or had an EC_{50} in the high nanomolar to low micromolar range.

Table 10

Receptor	Sequence	SEQ ID NO	Species of origin	EC_{50} range
CEGPCR16	SPLGTMRF-NH_2	99	*C. elegans flp3*	
(SEQ ID	SAEPFGTMRF-NH_2	97	*C. elegans flp3*	
NO:36)	SADDSAPFGTMRF-NH_2	96	*C. elegans flp3*	
	ASEDALFGTMRF-NH_2	77	*C. elegans flp3*	
	EDGNAPFGTMRF-NH_2	82	*C. elegans flp3*	~ 60 - 300 nM
	EAEEPLGTMRF-NH_2	81	*C. elegans flp3*	
	GMPGVLRF-NH_2 (AF20)	87	*A. suum*	
	SVPGVLRF-NH_2	100	*C. elegans flp18*	
CEGPCR16	KSVPGVLRF-NH_2	92	*C. elegans flp18*	
(SEQ ID	DVPGVLRF-NH_2	80	*C. elegans flp18*	
NO:36)	SEVPGVLRF-NH_2	98	*C. elegans flp18*	
	DFDGAMPGVLRF-NH_2	120	*C. elegans flp18*	~ 0.7 - 4.6 μM
	EIPGVLRF-NH_2	121	*C. elegans flp18*	
	AVPGVLRF-NH_2 (AF3)	79	*A. suum*	
	GDVPGVLRF-NH_2 (AF4)	84	*A. suum*	

The CEGPCR12h receptor, and the related splice variant isoforms CEGPCR12c, CEGPCR12u, and CEGPCR12v were expressed in CHO cells as described for CEGPRC3 and used in [^{35}S]GTPγS assays to screen for peptide ligands. The results presented in Table 11, specifically presented for CEGPCR12h, apply to all CEGPCR12 isoforms. As shown therein, the most active peptide ligand was schistoFLRFamide and its N-terminally iodinated derivative, exhibiting EC_{50} values of 330 nM and 478 nM, respectively. Candidate peptide ligands were tested at 5 μM.

Table 11

Receptor	Peptide sequence	SEQ ID NO	Species of origin	EC_{50} range
CEGPCR12h (SEQ ID NO:8)	PDVDHVFLRF-NH_2 (SchistoFLRF-NH_2)	94	Locust	~300-500 nM
	[I]Y^0PDVDHVFLRF-NH_2	103	Locust (modified)	
	FDDYGHLRF-NH_2	180	Drosophila	
CEGPCR12h (SEQ ID NO:8)	pQDVDHVFLRF-NH_2 (leucomyosuppressin)+	95	Locust Drosophila	~1.5-3.0 μM
	TDVDHVFLRF-NH_2 (dromyosuppressin)	101	*C. elegans*	
	GGPQGPLRF-NH_2	85		
CEGPCR12h (SEQ ID NO:8)	AVPGVLRF-NH_2 (AF3)	79	*A. suum*	Peptides active at 5 μM
	GDVPGVLRF-NH_2 (AF4)	84	*A. suum*	
	GMPGVLRF-NH_2	87	*A. suum*	
	GLGPRPLRF-NH_2 (AF9)	86	*C. elegans, A. suum*	
	DVPGVLRF-NH_2	80	*C. elegans/flp18*	
	KSVPGVLRF-NH_2	92	*C. elegans/flp18*	
	SVPGVLRF-NH_2	100	*C. elegans/flp18*	

+ pQ denotes pyroglutamic acid residue.

The CEGPCR19.1 and CEGPCR19.2 receptors, expressed in CHO cells, were subjected to screens for peptide ligands (present at 10 μM each) using a calcium mobilization assay (FLIPR). This calcium flux assay provides an alternative to the use of other screening assays disclosed herein. The assay, known in the art, was performed on transiently transfected CHO cells that had been incubated for 24 hours at 37°C immediately following transfection, followed by a 48-hour period at 28°C. Cells were then loaded with 4 M Fluo-3 and 2.5 mM probenecid for calcium signaling (FLIPR) analysis. The results shown in Table 12 indicate that CEGPCR19.1 was activated by Flp18 peptides and some structurally similar *A. suum* peptides (identical C-terminal sequences of PGVLRF-NH_2). EC_{50} values for CEGPCR19.1 binding of Flp18 peptides, and AF3 and AF4 peptides, ranged from 20-100 nM (similar EC_{50} values were observed for CEGPCR19.2, described below). The results support the conclusion that CEGPCR19.1 is either an isoform of, or closely related to, the CEGPCR4 receptor

because both CEGPCR19.1 and CEGPCR4 were activated by the same Flp18 peptides as well as by *A. suum* peptides sharing a high degree of similarity with those Flp18 peptides. CEGPCR19.2, which is a splice variant of CEGPCR19.1, was activated (calcium assay) by the same peptides that activated CEGPCR19.1, as well as by several additional *A. suum* peptides not identified as activators of CEGPCR19.2 (SEQ ID NOS:91, 99, 131, 133, 150 and 151), as shown in Table 13.

Table 12

Receptor	Peptide sequence	SEQ ID NO	Species of origin	EC_{50} range
CEGPCR19.1 (SEQ ID NO:107)	DVPGVLRF-NH_2	80	*C. elegans/flp18*	*~30 - 100 nM*
	KSVPGVLRF-NH_2	92	*C. elegans/flp18*	
	SEVPGVLRF-NH_2	98	*C. elegans/flp18*	
	SVPGVLRF-NH_2	100	*C. elegans/flp18*	
	DFDGAMPGVLRF-NH_2	120	*C. elegans/flp18*	
	EIPGVLRF-NH_2	121	*C. elegans/flp18*	
	AVPGVLRF-NH_2 (AF3)	79	*A. suum*	
	GDVPGVLRF-NH_2 (AF4)	84	*A. suum*	
CEGPCR19.1 (SEQ ID NO:107)	ARGPQLRLRF-NH_2	136	*L. decemlineata*	~200 -1500 nM
	GMPGVLRF-NH_2	87	*A. suum*	
	GLGPRPLRF-NH_2 (AF9)	86	*A. suum, C. elegans/flp21*	

Table 13

Receptor	Peptide sequence	SEQ ID NO	Species of origin	EC_{50} range
CEGPCR19.2 (SEQ ID NO:105)	DVPGVLRF-NH_2	80	*C. elegans/flp18*	~20-70 nM
	KSVPGVLRF-NH_2	92	*C. elegans/flp18*	
	SEVPGVLRF-NH_2	98	*C. elegans/flp18*	
	SVPGVLRF-NH_2	100	*C. elegans/flp18*	
	DFDGAMPGVLRF-NH_2	120	*C. elegans/flp18*	
	EIPGVLRF-NH_2	121	*C. elegans/flp18*	
	AVPGVLRF-NH_2 (AF3)	79	*A. suum*	
	GDVPGVLRF-NH_2 (AF4)	84	*A. suum*	
CEGPCR19.2 (SEQ ID	GLGPRPLRF-NH_2 (AF9)	86	*A. suum, C.elegans/flp21*	

Receptor	Peptide sequence	SEQ ID NO	Species of origin	EC_{50} range
NO:105)	AEGLSSPLIRF-NH_2	150	*A. suum*	~0.6 – 5 μM
	FDRDFMHF-NH_2	151	*A. suum*	
	AGPRFIRF-NH_2	133	*A. suum*	
	GMPGVLRF-NH_2	87	*A. suum*	
	ARGPQLRLRF-NH_2	136	*L. decemlineata*	
CEGPCR19.	KPNFIRF-NH_2 (PF4)	165	*P. redivivus*	
2 (SEQ ID	KPNFLRF-NH_2	91	*C. elegans/flp1*	peptides active
NO:105)	SQPNFLRF-NH_2	166	*C. elegans/flp1*	at 10 μM
	SPLGTMRF-NH_2	99	*C. elegans/flp3*	
	DPSFLRF-NH_2	131	*Manduca*	
	LQPNFLRF-NH_2	167	*H. contortus*	

G. MAP Kinase Activity Assay

Evaluation of MAP Kinase activity in cells expressing a GPCR-like receptor provide another assay to identify modulators of GPCR-like polypeptide activity. [See, *e.g.*, Lajiness *et al., Journal of Pharmacology and Experimental Therapeutics, 267(3)*: 1573-81 (1993); and Boulton *et al., Cell, 65*: 663-75 (1991).]

In one embodiment, CHO cells stably transfected with a GPCR-like receptor are seeded into 6-well plates at a density of 70,000 cells/well 48 hours prior to the assay. During this time, the cells are cultured at 37°C in DMEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 Units/ml penicillin and 10 μg/ml streptomycin. The cells are starved for serum for 1-2 hours prior to the addition of stimulants.

For the assay, the cells are treated with medium alone or medium containing a putative agonist, or the positive control phorbol myristoyl acetate (*i.e.*, PMA), and the cells are incubated at 37°C for varying times. To stop the reaction, the plates are placed on ice, the medium is aspirated, and the cells are rinsed with 1 ml of ice-cold PBS containing 1 mM EDTA. Thereafter, 200 μl cell lysis buffer (12.5 mM MOPS, pH 7.3, 12.5 mM β-glycerophosphate, 7.5 mM $MgCl_2$, 0.5 mM EGTA, 0.5 mM sodium vanadate, 1 mM benzamidine, 1 mM dithiothreitol, 10 μg/ml leupeptin, 10 μg/ml aprotinin, 2 μg/ml pepstatin A, and 1 μM okadaic acid) is added to the cells. The cells

are scraped from the plates and homogenized by 10 passages through a 23 3/4 needle. The cytosol fraction is then prepared by centrifugation at 53,000 rpm for 15 minutes.

Aliquots (5-10 µl containing 1-5 µg protein) of cytosols are mixed with 1 mM MAPK Substrate Peptide (APRTPGGRR, Upstate Biotechnology, Inc., N.Y.) and 50 µM [γ-^{32}P]ATP, (NEN, 3000 Ci/mmol) diluted to a final specific activity of ~2000 cpm/pmol in a total volume of 25 µl. The samples are incubated for 5 minutes at 30°C, and reactions are stopped by spotting 20 µl on 2 cm^2 of Whatman P81 phosphocellulose paper. The filter squares are washed in four changes of 1% H_3PO_4, and the squares are counted by liquid scintillation spectroscopy. Equivalent cytosolic extracts are incubated without MAPK substrate peptide, and the cpm from these samples are subtracted from the matched samples with the substrate peptide. The cytosolic extract from each well is used as a separate point. Protein concentrations are determined by a dye binding protein assay (Bio-Rad). Agonist activation of the receptor is expected to result in up to a five-fold increase in MAPK enzyme activity. This increase is blocked by antagonists.

H. [^{3}H]Arachidonic Acid Release

The activation of GPCRs has been observed to potentiate arachidonic acid release in cells, providing yet another useful assay for modulators of GPCR-like receptor activity. [See, *e.g.*, Kanterman *et al.*, *Molecular Pharmacology, 39*: 364-9 (1991).] For example, CHO cells that are stably transfected with a GPCR-like receptor expression construct are plated in 24-well plates at a density of 15,000 cells/well and grown in DMEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 Units/ml penicillin and 10 µg/ml streptomycin for 48 hours at 37°C before use. Cells of each well are labeled by incubation with [^{3}H]-arachidonic acid (Amersham Corp., 210 Ci/mmol) at 0.5 µCi/ml in 1 ml DMEM supplemented with 10 mM HEPES, pH 7.5, and 0.5% fatty-acid-free bovine serum albumin for 2 hours at 37°C. The cells are then washed twice with 1 ml of the same buffer.

Candidate modulator compounds are added in 1 ml of the same buffer, either alone or containing 10 µM ATP as a positive control and the cells are incubated at 37°C for 30 minutes. Buffer alone and mock-transfected cells are used as controls. Samples (0.5 ml) from each well are counted by liquid scintillation spectroscopy. Agonists which

activate the receptor will lead to potentiation of the ATP-stimulated release of [^{3}H]-arachidonic acid. This potentiation is blocked by antagonists.

I. Extracellular Acidification Rate

In yet another assay, the effects of putative modulators of GPCR-like receptor activity are assayed by monitoring extracellular changes in pH induced by the putative modulators. [See, *e.g.*, Dunlop *et al., Journal of Pharmacological and Toxicological Methods, 40(1)*: 47-55 (1998).]

CHO cells transfected with a GPCR-like receptor expression construct are seeded into 12 mm capsule cups (Molecular Devices Corp.) at 4 x 10^5 cells/cup in DMEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 10 Units/ml penicillin, and 10 µg/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO_2 for 24 hours.

Extracellular acidification rates are measured using a Cytosensor microphysiometer (Molecular Devices Corp.). The capsule cups are loaded into the sensor chambers of the microphysiometer and the chambers are perfused with running buffer (bicarbonate free DMEM supplemented with 4 mM L-glutamine, 10 Units/ml penicillin, 10 µg/ml streptomycin, 26 mM NaCl) at a flow rate of 100 µl/minute. Agonists or other agents are diluted into the running buffer and perfused through a second fluid path. During each 60-second pump cycle, the pump is run for 38 seconds and is off for the remaining 22 seconds. The pH of the running buffer in the sensor chamber is recorded during the cycle from 43-58 seconds, and the pump is re-started at 60 seconds to start the next cycle. The rate of acidification of the running buffer during the recording time is calculated by the Cytosoft program. Changes in the rates of acidification are calculated by subtracting the baseline value (the average of four rate measurements immediately before addition of modulator candidates) from the highest rate measurement obtained after addition of a modulator candidate. The selected instrument detects 61 mV/pH unit. Modulators that act as agonists at the receptor result in an increase in the rate of extracellular acidification as compared to the rate in the absence of agonist. This response is blocked by modulators which act as antagonists at the receptor.

Example 6

To avoid impediments imposed by potential artifacts of PCR-based cloning, alternative approaches to the cloning of neuropeptide (*e.g.*, FaRP) receptors was developed. By increasing the versatility of the cloning approach, FaRP receptors having desired relationships (*e.g.*, particular levels of primary, or structural, sequence similarity or identity, common activities, or related sources) could be targeted. For example, targeted neuropeptide (*e.g.*, FaRP) receptors might exhibit 70%, 75%, 80%, 85%, 90%, 95%, 99%, 99.9%, or 99.99% similarity; also, the targeted receptors might be selected from invertebrates, from insects, from flat- and round-worms, from particular round-worm genera, or from particular round-worm species, sub-species or variants.

The identification of several authentic neuropeptide (FaRP) receptors also makes possible additional bioinformatics strategies. Comparison of confirmed receptors with the *C. elegans* database should aid in the identification of additional neuropeptide receptors, including FaRP receptors. Likewise, the identification of the *C. elegans* neuropeptide receptors make possible a similar approach using other databases, including Drosophila and human databases.

The sequences of a total of 20 cloned receptors are listed, including all of the receptors exhibiting high levels of similarity to the confirmed neuroepeptide (FaRP) receptors. Finally, the results of bioinformatic analyses using the current databases and biochemical information are presented.

Bioinformatic selection of receptors

Later releases of the Wormpep database were searched in a manner similar to that described above for the initial identification of candidate receptors. Wormpep releases through Wormpep 23 were downloaded from the Sanger Centre web site. The database was searched using BLAST 2.0, Altschul *et al.*, Nucl. Acids Res. 25(17):3389-402 (1997), with known *C. elegans* neuropeptide (*e.g.*, FaRP) receptors (CEGPCR3, CEGPCR4, CEGPCR5, CEGPCR7, CEGPCR12c, CEGPCR12h, CEGPCR12u, CEGPCR12v, and CEGPCR16). A composite BLAST score was assembled from these searches by calculating the product of the "e" values from all of the searches for a given receptor. Release 118 of GenPept was used for general searches against public sequences and as a source of sequences used for the phylogenetic analyses. Alignments

and phylogenetic relationships were prepared using the AlignX (ClustalW, Thompson *et al.*, Nucl. Acids Res. 22(22):4673-80 (1994)) alignment program (Vector NTI suite, Informax, N. Bethesda, MD).

Cloning of receptors

Receptors were cloned using PCR techniques. Initially, the predicted sequence was used to design PCR primers. The 5' PCR primer incorporated the *C. elegans* initiator ATG within an optimized 5' untranslated region (Kozak, M.., Nucl. Acids Res. 15(20):8125-48 (1987). The 3' PCR primer usually contained a restriction enzyme site not contained within the receptor sequence to simplify later production of cRNA for injection studies.

If the initial approach using primers designed to the predicted sequence did not readily produce a PCR product, or if the predicted sequence was obviously in error (as judged by alignment with other GPCRs), then either an anchored PCR approach (used successfully for CEGPCR4 and CEGPCR22), or a rational primer design approach (see below) was followed to obtain an intact clone.

Typically, a PCR product was obtained using either XL rTth polymerase (Perkin-Elmer Applied Biosystems, Foster City, CA) or the Expand HF mix (Boehringer Mannheim, Indianapolis, IN). PCR primer sequences are presented in the sequence listing, with "f" and "r" suffixes added to a given GPCR clone designation to unambiguously refer to forward and reverse primer pairs. The DNA template for the reaction was either a first strand cDNA synthesis reaction (using random hexamers to prime synthesis from a preparation of total *C. elegans* RNA) or a *C. elegans* cDNA library in pBluescript, constructed by performing an *in vivo* excision reaction on a cDNA commercial library constructed in a Uni-ZAP XR lambda vector (Stratagene, La Jolla, CA, catalog number 937006). PCR products were typically gel-purified using QIAquick columns (Qiagen, Valencia, CA) and either cloned directly into pCR3.1 (Invitrogen, Carlsbad, CA), or indirectly through an initial cloning step into pCR2.1. Plasmids containing the appropriate inserts were prepared by a mini-prep DNA isolation procedure (Qiagen) and were sequenced using BigDye dye terminator chemistry (PE Applied Biosystems, Foster City, CA) on an ABI 377 automated DNA sequencer, as described above. Complete sequences were assembled using Sequencher software (Version 3.0,

GeneCodes, Ann Arbor, MI) and compared against the expected sequence predicted from the Wormpep database.

Cloning CEGPCR2

CEGPCR2 (SEQ ID NO:176) was cloned using PCR. The template for the reaction was a plasmid library of *C. elegans* cDNA isolated from an *in vivo* excision of a commercially available library (constructed in a Uni-ZAP XR vector; Stratagene, La Jolla, CA) according to the manufacturer's instructions. The PCR reaction was performed using the Expand HF kit (Boehringer Mannheim, Indianapolis, IN) and contained: 0.5μM each primer, 150 nM dNTPs, 0.5mM $MgCl_2$, and 1.4 units Expand High Fidelity enzyme mix. The PCR reaction was carried out in a Stratagene Robocycler, using the following conditions: 30 cycles at 95°C for 30 seconds, 50°C for 30 seconds, 72°C for 2 minutes, followed by 72°C for 5 minutes. A TA cloning kit (Invitrogen, Carlsbad, CA) was used to clone the gel-purified PCR product into pCR2.1. The ligation mixture was transformed into *E. coli* DH5α, and transformed cells were identified using a colony PCR method to screen transformants for the presence of the appropriate insert. Purified plasmid DNA was obtained using a mini-prep procedure (Qiagen Valencia, CA). Plasmids were sequenced using dye terminator chemistry (PE Applied Biosystems, Foster City, CA) and an automated sequencer (ABI 377). A clone with an intact open reading frame was obtained and subcloned into pCDNA3.1(+) using *Eco*RI. *Xho*I was used to determine the orientation of insert fragments.

Cloning CEGPCR21

CEGPCR21 was cloned as described above in the context of CEGPCR2, using the same template, PCR conditions, and cloning strategy. *Hin*dIII was used to determine the orientation of inserts.

Cloning CEGPCR23

The *C. elegans* receptor CEGPCR23 was cloned using the *C. elegans* cDNA plasmid library as template. The PCR reaction was set up using 0.5μM each primer, 200 μM dNTPs, 1.15mM $Mg(OAc)_2$, and 4 units rTth DNA polymerase, XL. The PCR reactions were performed in a Perkin-Elmer 9600 Thermal Cycler using the following program: 25 cycles at 94°C for 30 seconds, 50°C for 10 minutes, and 72°C for 10 minutes. The resulting PCR product was gel-purified, cloned into pCR2.1, and

propagated in *E. coli* DH5α. Plasmid DNA was isolated and verified by sequencing. An intact clone was selected and subcloned into pCDNA3.1(+) (Invitrogen) using *Bst*XI. Subclones with inserts in the correct orientation were identified using *Bam*HI/*Bgl*II.

Cloning CEGPCR22

An anchored PCR approach was used to obtain the correct ends for CEGPCR22. Internal primers were designed to the exon containing 7TM. PCR experiments using these internal primers with the initial flanking primers confirmed that the 3' PCR primer was nonfunctional. Using the internal primer (SCN162 – 5'-ACG TTT AAG AGC TCT CAA ATC CCA T-3'; SEQ ID NO:171) and T7/M13 reverse vector primers, along with the *C. elegans* cDNA plasmid library, a PCR product was successfully obtained that corresponded to the correct 3' end. Using that information, a new 3' PCR primer was designed and used in conjunction with the initial 5' PCR primer. CEGPCR22 was then cloned in a similar manner to CEGPCR23. Template and PCR conditions were identical to CEGPCR23, except that the PCR program included 30 cycles at 94°C for 15 seconds, 55°C for 5 minutes, and 72°C for 10 minutes. The resulting PCR product was directly cloned into pCR3.1.

Cloning CEGPCR4

An anchored PCR approach was used to obtain the correct ends for CEGPCR4. Internal primers were synthesized. PCR experiments using these internal primers with the initial flanking primers confirmed that the 3' end of the cDNA was incorrectly predicted. Using an internal primer (SCN160 - 5'-CCAGAGCTCATCAAA-ACTCAAGAAT-3'; SEQ ID NO:172) and T7/M13 reverse vector primers, along with the *C. elegans* cDNA plasmid library, a PCR product was successfully obtained that corresponded to the correct 3' end, extending beyond the stop codon. Using that information, a new 3' PCR primer (SCN189, 5'-TTACAATTTAAAACTAGGTGCTTCT-3'; SEQ ID NO:174) was designed and used in conjunction with the initial 5' PCR primer. CEGPCR4 was then cloned using identical conditions to CEGPCR23, except that the template was derived from a first strand cDNA synthesis of *C. elegans* mRNA and the PCR program included 25 cycles at 94°C for 15 seconds, 60°C for 5 minutes, and 72°C for 10 minutes. The PCR product was cloned

initially in pCR2.1, and subcloned into pCDNA3.1(+) using *Eco*RI, and clones with the correct orientation were found using digestion with *Sac*I.

Cloning CEGPCR5

CEGPCR5 was cloned using the same PCR reaction components as CEGPCR22 and first strand cDNA as template. PCR conditions were: 35 cycles at 94°C for 30 seconds, 60°C for 5 minutes, and 72°C for 10 minutes. The reactions were carried out in a Perkin-Elmer 9600 thermal cycler. The resulting PCR product was cloned directly into pCR3.1 and transformed into *E. coli* DH5α. Clones were identified by colony PCR. Plasmids were isolated, purified, and sequenced as described above.

Cloning CEGPCR19.1

The short splice variant of clone CEGPCR19 was cloned using conditions identical to those described for CEGPCR5, with the exception of a 50°C hybridization step for the PCR reaction.

Cloning CEGPCR19.2

The long splice variant of clone CEGPCR19 was cloned using conditions identical to those described for CEGPCR5, with the exception of a 55°C hybridization step for the PCR reaction.

Cloning CEGPCR1a and CEGPCR1f

The cloning of the two splice isoforms of CEGPCR1 were described in the examples above. Those initial clones were based on the Wormpep 13-predicted sequence. Although the receptors were successfully amplified and cloned, alignment of the sequences suggested that other variant clones might also exist that contained an initiator ATG in a more conventional location than within the first transmembrane region. Subsequent releases of the database included a longer AC7.1 sequence corresponding to CEGPCR1 (but still not the splice variants). Using that new sequence from the Wormpep database, a new primer (SCN-199, 5'-GCCGCCATGAACTTTTCGGCCACCGATTCGA-3'; SEQ ID NO:175) was designed and additional clones were isolated (using the identical procedures and 3' PCR primer described in the preceding examples for CEGPCR1).

Cloning CEGPCR12c, CEGPCR12h, and CEGPCR12u

Similar to CEGPCR1, alignment of the initially amplified clones for CEGPCR12 suggested that additional variants might exist. In the case of CEGPCR12, it was anticipated that variants encoding polypeptides having additional N-terminal sequence would be identified. Using the rational primer design approach, PCR primers were designed for several potential exons. Primer DEL-1850 (5'- GCCGCCATGT-CGAATGATCTCGTGCCTTCAG-3', SEQ ID NO:173) and the original 3' PCR primer (SEQ ID NO:70; using PCR conditions as described for the initial amplification of CEGPCR12) were then used to amplify the longer version. The PCR product was cloned directly into PCR3.1 as described above. Intact versions for CEGPCR12c, CEGPCR12h, and CEGPCR12u were obtained.

The *C. elegans* Wormpep 23 database was searched using BLAST 2.0 (Altschul, *et al.*, Nucl. Acids Res. 25(17):3389-402 (1997)), with the six currently known *C. elegans* neuropeptide (*e.g.*, FaRP) receptors (CEGPCR3, CEGPCR4, CEGPCR5, CEGPCR7, CEGPCR12c, CEGPCR12h, CEGPCR12u, CEGPCR12v, and CEGPCR16). Using the e-values from these searches, a composite BLAST score was calculated. Using this approach, a group of 11 receptors was discernable from the rest of the *C. elegans* GPCRs. These receptors include the above-six confirmed neuropeptide (FaRP) receptors, and CEGPCR13, CEGPCR17, CEGPCR14, CEGPCR19 (both CEGPCR19.1 and CEGPCR19.2), and CEGPCR11. An additional group of receptors fell into a second tier. This group included CEGPCR20, CEGPCR1 (Both CEGPCR1a and CEGPCR1f), CEGPCR15, CEGPCR18a, CEGPCR25 (SEQ ID NO:178), and CEGPCR3. BLAST comparisons against the GenBank database showed that these receptors were also likely peptide receptors, although they have not yet been matched to ligands.

An alignment of the receptors using CLUSTALW shows that the receptors exhibit the classic characteristics of the seven transmembrane family of receptors, including higher degrees of similarity through the transmembrane regions and very few absolutely conserved residues (for a review see Probst *et al.*, Cell Biol 11(1):1-20 (1992)). The only unusual feature for the receptors is a few atypical amino acids in the conserved 'DRY' motif immediately following 3TM, *e.g.*, 'HEF' in the CEGPCR1a and 1f sequences and 'DKF' for the CEGPCR18a and 18b sequences.

The *C. elegans* receptors can also be classified using a phylogenetic approach. Figure 1 shows an alignment of the *C. elegans* GPCRs described above resulting from a typical phylogenetic analysis. This analysis showed that ten of the eleven members of the above-mentioned family of receptors, including the neuropeptide receptors identified to date, fall into two distinct groups. The only other receptors that fall into these groups are the 'GRL105' receptor from *Lymnaea stagnalis* (an invertebrate NPY receptor) and the DmGPCR1 and DmGPCR2 receptors from Drosophila. We have recently confirmed that the DmGPCR1 receptor has a neuropeptide (*i.e.*, FaRP) ligand.

This analysis also showed that the second tier of *C. elegans* receptors from the BLAST analysis tends to fall into classes with other known neuropeptide receptors. CEGPCR1 and CEGPCR24 fall into an interesting group with the *L. stagnalis* lymnokinin receptor (Cox *et al.*, J. Neurosci. 17(4):1197-205 (1997)), the *L. stagnalis* cardioexcitatory peptide (an RFamide ligand) receptor (Tensen *et al.*, J. Neurosci. 18(23):9812-21 (1998)), the *Stomoxys calcitrans* (stable fly) 'tachykinin-like' receptor (Guerrero, F.D., Annals N.Y. Acad. Sci. 814:310-1 (1997)), a receptor for a leucokinin-like peptide from cattle tick (Holmes, *et al.*, Insect Mol Biol (2000; in press)) and the DmGPCR6 receptor [Li, *et al.*, J. Biol. Chem. 267(1):9-12 (1992). We have recently identified this last receptor as having a neuropeptide (*i.e.*, FaRP) ligand. This group is separate, and contains only invertebrate receptors, but is most closely related to the vertebrate family of neurokinin (NK-1,2,3) receptors. The *C. elegans* receptors CEGPCR6 and CEGPCR15 are in a group distinct from, but related to, the vertebrate neurokinin receptors. A cluster containing a variety of cholecystokinin receptors contains the *C. elegans* receptor Y39A3B.5. The receptor CEGPCR20 falls into a group containing the Drosophila receptor for an allatostatin-related peptide (Birgul *et al.*, EMBO Journal 18(21):5892-900 (1999)) and DmGPCR4, both of which recognize allatostatin peptide ligands. CEGPCR18 variants are included in a group containing several vertebrate orphan receptors, most closely related to the clusters containing bombesin receptors, galanin receptors, gastrin releasing peptide receptors and the above mentioned allatostatin receptors. CEGPCR8 (and CEGPCR21) cannot obviously be linked to any particular group of receptors, being most closely related to a broad cluster containing a variety of receptors, including some falling outside the domain of those

receptors recognizing peptide ligands (*e.g.*, serotonin receptors, opsins). The receptors CEGPCR9, CEGPCR23 and CEGPCR22 are all listed in a small group apart from a group containing a variety of receptors with peptide ligands, including somatostatin, galanin, and bombesin receptors. BLAST results are slightly in contrast with these results for CEGPCR9, listing the highest scoring hits for this receptor as NPY-type receptors. BLAST results indicate the five highest scoring matches to CEGPCR22 and CEGPCR23 are all other *C. elegans* receptors (Y116A8B.5, C43C3.2, T02D1.6), but immediately following those are a series of vertebrate somatostatin and kappa-opioid receptors.

The receptors were cloned using several strategies, all based on the use of PCR. The first strategy involved the simple design of PCR primers using the predicted cDNA sequence of the GPCR. In cases where a PCR product was not obtained using primers designed on the basis of the predicted sequence, an alternative approach was used to design new PCR primers. The failure in the initial PCR experiments was typically due to the inability of the gene prediction software to accurately predict the ends of the GPCR sequences. Alignment of the predicted amino acid sequences for the GPCRs often revealed the errors, *e.g.*, one of the ends of the encoded protein sequences would be much longer that any of the other aligned sequences. In some cases the alignment showed that a product that did amplify by PCR had been inaccurately predicted, *e.g.*, showing an initiator ATG that was within the predicted 1TM region. The incorrectly predicted end could usually be pinpointed by the use of internal PCR primers, located within a conserved and unspliced region, in conjunction with the original flanking PCR primers. Using this information, new PCR primers were designed. Taking into account the possible open reading frames, potential mRNA processing sites, and proximity to the presumed correct sequence, a series of alternative PCR primers were synthesized and tested. This rational primer design approach, although labor-intensive, was extremely successful, resulting in a high rate of successful PCR amplification.

Based on comparisons of cloned cDNAs to predicted sequences, 11 of 22 GPCR-like receptor sequences selected from Wormpep were incorrectly predicted. As would be expected, most of these errors have occurred in the prediction of the N- and C-termini of the encoded proteins. Although an anchored PCR approach has been

successful in obtaining the ends of an incorrectly predicted GPCR-like receptor in a couple of cases, the rational primer design approach, although laborious, has had a much higher success rate.

Phylogenetic analyses can be used to speculate on possible functions for the receptors. It should be noted that there are numerous methods and variables that can be used to build phylogenetic relationships. Although the results described herein are based on one method Thompson *et al.*, Nucl. Acids Res. 22(22):4673-80 (1994)), several methods and permutations of phylogenetic analyses were investigated, with all methods giving very similar results, especially with regards to the overall relationships observed for the confirmed neuropeptide (*e.g.*, FaRP) receptors.

At this point in time, a total of six *C. elegans* receptors and two Drosophila receptors have been matched to neuropeptide (FaRP) ligands. Phylogenetic analyses of the *C. elegans* and Drosophila receptors, along with all of the peptidergic receptors in the public databases, show that all of the *C. elegans* FaRP receptors may be grouped together. This group contains an additional five *C. elegans* receptors (CEGPCR13, CEGPCR14, CEGPCR17, and CEGPCR19.1 and CEGPCR19.2) that are expected to have FaRP ligands. The only other receptors within this group are two Drosophila receptors, one of which has been confirmed to have a FaRP ligand (DmGPCR1), and a single receptor from *L. stagnalis* (GRL105) (Tensen *et al.*, J. Neurosci. 18(23):9812-21 (1998)). The GRL105 receptor has been linked to a *L. stagnalis* neuropeptide Y homologue (an RFamide). The phylogenetic analysis would predict that it may also have a FaRP ligand. The grouping of these invertebrate receptors in a category that contains no known vertebrate receptors is consistent with FaRP receptors providing a relatively invertebrate-specific target.

The other *C. elegans* receptor that received a high score in the BLAST analyses was CEGPCR11. This receptor falls into a group that is highly related to the group containing confirmed FaRP receptors, described above, and includes the dog and human orexin receptors. Slight alterations in the conditions used in the analysis would result in CEGPCR11 being grouped with the above receptors.

Another group of receptors that is of interest is derived from *C. elegans* neuropeptide receptors. This group includes the *C. elegans* receptors CEGPCR24a,

CEGPCR24b, and CEGPCR1. The recent identification of a FaRP ligand for the Drosophila receptor DmGPCR6, also located within this group of receptors, increases the interest in these *C. elegans* receptors.

Two clones (CEGPCR6 and CEGPCR20) included in the table have yet to be amplified. These clones do not PCR-amplify using primers designed from the predicted sequences in Wormpep. It is very likely that one, or both, of the termini of these clones is incorrectly predicted. At the time of the receptor cloning, these receptors were not classified in the group of receptors that contained the first verified FaRP receptors. CEGPCR20 groups with a family of receptors that include the vertebrate galanin receptors, the Drosophila allatostatin-like peptide receptor (Birgul *et al.*, EMBO Journal 18(21):5892-900 (1999)), the Drosophila 'galanin' receptor Lenz *et al.*, BBRC 269(1):91-96 (2000) (ligand assigned by homology, the Drosophila receptor DmGPCR4, which has recently been confirmed to be an additional Drosophila allatostatin receptor, and a series of four orphan *L. stagnalis* receptors. Based on this information, is expected that CEGPCR20 recognizes an allatostatin-like peptide ligand. CEGPCR6 falls into a grouping with CEGPCR1a, CEGPCR1f, CEGPCR15, and two Drosophila receptors. With the recent discovery that one of the Drosophila receptors recognizes a FaRP ligand, there is a renewed interest in this group of receptors.

Based on analyses to date, the invertebrate GPCR-like receptors do not appear to have highly similar sequences in other organisms, such as vertebrates and plants. However, receptors bearing lower levels of similarity, *e.g.*, 55%, 60%, 65%, 70%, 75%, 80%, 85%, and preferably 90%, 95%, 98%, 99% and more preferably 99.5% similarity to a GPCR-like receptor amino acid sequence disclosed herein are also contemplated by the invention. Analogously, the invention comprehends receptor-encoding polynucleotides exhibiting 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, and preferably 95%, 98%, 99% and more preferably 99.5% similarity to a polynucleotide disclosed herein. Similarities can be determined using any of a variety of algorithms known in the art, with the BLAST algorithm implemented at the GenBank website (http://www.ncbi.nlm.nih.gov/BLAST/), using default parameters, being preferred. These receptor sequences are anticipated to be useful in a variety of contexts. For example, it is expected that vertebrate, and more particularly mammalian, receptor sequences showing some similarity to the invertebrate

GPCR-like receptors will be useful in diagnosing, and treating, a variety of neurological ailments or conditions.

The following clones (deposit numbers indicated parenthetically) were deposited with the ARS Patent Culture Collection, 1815 North University Street, Peoria, Illinois 61604 on November 3, 2000: CEGPCR1a (UC20129 NRRL B-30361), CEGPCR1f (UC20130 NRRL B-30362), CEGPCR3 (UC20132 NRRL B-30364), CEGPCR4 (UC20133 NRRL B-30365), CEGPCR5 (UC20134 NRRL B-30366), CEGPCR7 (UC20135 NRRL B-30367), CEGPCR8 (UC20136 NRRL B_30368), CEGPCR9 (UC20137 NRRL B-30369), CEGPCR12c (UC20138 NRRL B-30370), CEGPCR12h (UC20139 NRRL B-30371), CEGPCR12u (UC20140 NRRL B-30372). On November 10, 2000, the following clones were deposited with the ARS Patent Culture Collection: CEGPCR11 (UC20141 B-30381), CEGPCR13 (UC20142 B-30382), CEGPCR14 (UC20143 B-30383), CEGPCR15 (UC20144 B-30384), CEGPCR16 (UC20145 B-30385), CEGPCR17 (UC20146 B-30386), CEGPCR18a (UC20147 B-30387), CEGPCR19.2 (UC20148 B-30388), CEGPCR19.1 (UC20149 B-30389), CEGPCR21 (UC20150 B-30390), CEGPCR22 (UC20151 B-30391), and CEGPCR23 (UC20152 B-30392).

Numerous modifications and variations in the invention as set forth in the above illustrative examples are expected to occur to those skilled in the art. Consequently only such limitations as appear in the appended claims should be placed on the invention.